Molecular mechanisms of endosymbiosis between cnidarian animals and symbiotic algae

博士論文

Molecular mechanisms of endosymbiosis between cnidarian animals and symbiotic algae
(サンゴ共生藻と刺胞動物との細胞内共生の分子メカニズム)

令和元年度

東北大学大学院生命科学研究科
生態発生適応科学専攻 進化生物分野

石井　悠

Abstract

Symbiosis is a living state in which individuals of multiple species coexist interactively and sympatrically. Symbiotic relationship is considered to contribute to the enhancement of the survival fitness or the adaptation to new environment by the individuals involved in the system. Some cnidarians such as corals and sea anemones live with Symbiodiniaceae dinoflagellate algae by harboring the algae in their own endodermal cells (endosymbiosis). It is presumed that the cnidarians provide a stable habitat and inorganic nitrogen compound and carbon dioxide to the algae as "host", while the algae provide photosynthetic products to the host as "symbiont". This symbiotic relationship is thought to be advantageous in oligotrophic tropical and subtropical oceans, but can be vulnerable to environmental changes. Recent studies have revealed huge impact of elevated seawater temperature on the collapse of their symbiosis (Hughes *et al.* 2017). Nevertheless, the molecular mechanism of the symbiosis between the cnidarians and the dinoflagellate algae are still unknown. To gain insights into the molecular interaction in the symbiosis, here I analyzed the candidate genes involving in symbiotic state changes, i.e. changes between symbiotic and apo-symbiotic (symbiont free) states (Chapter 1) and established mutant symbiont algal strains as a way to test their functions (Chapter 2). As experiment materials, I used the model symbiotic sea anemone *Exaiptasia diaphana* with its Symbiodiniaceae symbiont algae. *E. diaphana* has a substantial advantage in easiness to maintain in laboratory even if they are under the apo-symbiotic state and to manipulate the symbiotic state by experimentally switching between symbiotic and apo-symbiotic conditions. In Chapter 1, I performed transcriptomic analyses under multiple conditions using the symbiotic and apo-symbiotic *E. diaphana*. Comparative analysis of expression profiles under multiple conditions highlighted candidate genes potentially important in the symbiotic state transition under heat-induced bleaching. Many of these genes were functionally associated with carbohydrate and protein metabolisms in lysosomes. Symbiont algal genes differentially expressed *in hospite* encode proteins related to heat shock response, calcium signaling, organellar protein transport, and sugar metabolism. These results suggest that heat stress alters gene expression in both the hosts and symbionts. In particular, heat stress may affect the lysosome-mediated degradation and transportation of substrates such as carbohydrates through the membrane of symbiosome (phagosome-derived organelle harboring symbiont), which potentially might attenuate the stability of symbiosis and lead to bleaching-associated symbiotic state transition. Although gene introduction and manipulation techniques are necessary to directly test this hypothesis, any methods in both *E. diaphana* and Symbiodiniaceae had not been readily available. Therefore, in Chapter 2, I isolated Symbiodiniaceae mutants to use for transformation screening. *Breviolum* sp. (Symbiodiniaceae) was cultured in the presence of 5-fluoroorotic acid (5FOA), which inhibits the growth of wild type cells expressing *URA3* encoding orotidine-5'-monophosphate decarboxylase, and isolated spontaneous mutant cells that require uracil for growth. An obtained mutant cell line had a point mutation (splicing variation) in the *URA3* gene,

which was confirmed by sequence analyses and genetic complementation tests in the yeast. This mutant maintained a symbiotic relationship with *E. diaphana* in sea water containing uracil but didn't without uracil. This *URA3* mutant will be a useful tool for screening Symbiodiniaceae transformants, both ex and in *hospite*, as survival in the absence of uracil is possible only upon successful introduction of a functional *URA3* gene. In conclusion, these results have provided a foundation for gene function analysis for candidate genes in this thesis in future studies.

Acknowledgements

This thesis would not have been possible without the generous aid of so many people that I have come to meet over its course. I would like to give heartful thanks to my supervisor Prof. Masakado Kawata who provided helpful comments and suggestions. Special thanks also go to Dr. Shinichiro Maruyama who gave me invaluable comments and warm encouragements. Thanks should also be extended to the Kawata Lab Group, past and present, for their positivity, creative thinking and words of kindness along the way.

I would like to thank my co-authors. Thanks to Minagawa Lab Group, in particular to Prof. Jun Minagawa, Dr. Shunichi Takahashi, Dr. Konomi Kamada and Dr. Yusuke Aihara, who gives insightful comments and suggestions. A big thank you also to Ueno Lab Group, in particular to Prof. Naoto Ueno, Dr. Hiroki Takahashi and Dr. Takeshi Yamaguchi, who gives insightful suggestions and teach me a lot of technique for molecular experiments. Prof. Shuji Shigenobu and Dr. Katsushi Yamaguchi supported to get data. Dr. Natsumaro Kutsuna teach me how to analyze image data.

Thank you all.

Table of contents

General introduction

Interspecies relationship is diverse. Not only limited to the prey-predator relationship, some species take advantages of ecological or physiological characteristics of different species, exploitatively or cooperatively. Although its definition has been a debate, these interspecies relationships are generally called 'symbiosis'. In this thesis, symbiosis is defined as a living state in which individuals of multiple species coexist interactively and sympatrically. Symbiosis is considered to be an evolutionary force that enables the living organisms to enhance their survival fitness and to adapt to new environment via interaction with symbiotic partners. The dependency to symbiotic partner species is widely varied. Not a few species have developed indispensable relationships as they cannot live without their partner(s).

Some cnidarians such as corals and sea anemones live with Symbiodiniaceae dinoflagellate algae, formerly called 'zooxanthellae', which are harbored in their cells (endosymbiosis). The family Symbiodiniaceae are difficult to classify based on morphological characters and genetically classified into 9 clades, many of which are corresponding to genus-level classifications, using molecular phylogeny-based approaches (LaJeunesse *et al.* 2018). Normally a host animal is symbiotic with one or several clades of Symbiodiniaceae (Stat *et al.* 2006). The cnidarians provide stable habitat and nitrogen source/carbon dioxide to the algae as "host" while the algae provide photosynthetic products to the host as "symbiont". This relationship has been thought to be a key for their prosperities in oligotrophic tropical and subtropical oceans. Although they have developed interactions in intracellular level, actually their symbiosis is not necessarily crucial for a short term. Their relationship is plastic: it can be collapsed and re-established. This contrasts with static endosymbiosis with prokaryotes which gave births to mitochondria and chloroplast in the history of eukaryotic cell evolution. Symbiotic cnidarians take the dinoflagellate algae into the cells at the planula stage in their development or after the transformation to polyp(s), while some species acquire the symbiont at a very early embryo stage (Yamashita and Koike 2011). The algae are taken into the gastrovascular cavity from the mouth, and are taken into the endodermal cells by endocytosis. Conversely, the algae are also expelled out of mouse via the gastrovascular cavity. In addition, the algae can be eliminated within the cells by host phagocytosis. Nevertheless, many basic questions on symbiosis remains unanswered: e.g. how symbiosis is maintained, what switches between the symbiotic and apo-symbiotic (symbiont-free) states?

Environment changes have a huge impact on the symbiosis between the cnidarians and the dinoflagellate algae. Although the cnidarians can survive without symbionts in the short term, some species need to re-establish the symbiotic relationship for growth in the long run. Especially, elevated water temperature is one of major factors, which can cause the collapse of their symbiosis. When the collapse happens in whole body of the cnidarian, the body color becomes white due to loss of the algae's photosynthetic pigments or algae themselves, that is known as 'bleaching' (Takahashi *et al.* 2008). Coral reefs are now endangered since sea water temperature elevation takes place more frequently due to recent climate

changes, e.g. global warming. The coral reefs not only play a primary producer in oligotrophic oceans but also provide habitats for small marine organisms in the spaces structured by their complexed bodies, that is, so to speak, a key player to maintain biological diversity in tropical/subtropical oceans. Therefore, further understanding of precise mechanism of bleaching is of urgent need.

Whilst the environmental factors causing bleaching have been widely studied (Brown 1997), the molecular mechanism of the symbiosis between cnidarians and dinoflagellate algae are poorly understood. One major problem is difficulty in keeping/handling corals in laboratory tanks. In wild, it is also technically difficult to find out symbiotic and apo-symbiotic state colonies of the same genetic background and compare gene expression patterns between them grown under uncontrolled environments. Therefore, instead of using coral, a symbiotic sea anemone, *E. diaphana,* and its Symbiodiniaceae symbionts have been focused as a model experimental system for cnidarian-algal symbiosis (Baumgarten *et al.* 2015). *E. diaphana* has big advantages in easiness to keep in laboratory even if they are under apo-symbiotic state and to manipulate the symbiotic state by experimentally inducing the switch between symbiosis and apo-symbiosis. Genome sequencing has been completed in *E. diaphana* (Baumgarten *et al.* 2015) and 5 species of the family Symbiodiniaceae (Shoguchi *et al.* 2013; Lin *et al.* 2015; Aranda *et al.* 2016; Liu *et al.* 2018; Shoguchi *et al.* 2018). Previous transcriptomic and proteomic studies reported that in host sea anemones many metabolic pathways and cellular processes were affected by symbiosis based on the comparison between the symbiotic and apo-symbiotic states (Kuo *et al.* 2004; Rodriguez-Lanetty *et al.* 2006; Ganot *et al.* 2011; Lehnert *et al.* 2014; Oakley *et al.* 2016; Matthews *et al.* 2017). Meanwhile, a proteomic analysis using *E. diaphana* has identified a number of metabolic pathways located in cellular compartments (e.g. endoplasmic reticulum) as essential in the cellular response to heat stress (Oakley *et al.* 2017). Nevertheless, the interplay between symbiotic states and heat stress responses has not been well studied at the molecular level via genome-wide analyses.

The aim of this thesis is to identify the genes involving in the environmental responses, especially elevated temperature which can cause bleaching in wild coral reefs, in the sea anemone and the dinoflagellate algae. For the first approach, screening the candidate genes was conducted by using transcriptome analysis in a comparison of 4 different condition, symbiotic/apo-symbiotic states × normal/heat temperature of seawater (Chapter 1). Although heat stress can induce bleaching, the stress responses could be affected by the relationship between host and symbiont and may be different when hosts and symbionts live independent of the partners. The transcriptome analysis in this thesis was designed to detect symbiotic state-specific gene expression induced by heat stress. These genes are expected to be involved in heat induced bleaching process. To analyze the functions of these candidate genes and test hypotheses, gene manipulation would be a powerful tool. Recently, development of the gene introduction methods in host cnidarians is progressing; CRISPR/Cas9

system was reported to work in a coral (Cleves *et al.* 2018) and a microinjection method was developed to introduce gene fragments into *E. diaphana* (Jones *et al.* 2018). However, any reproducible and readily available gene manipulation methods for symbiont Symbiodiniaceae algae had not been established. So far, I have succeeded to establish a *Breviolum* (Symbiodiniaceae) auxotroph mutant strain 'T01', which can be a fundamental tool for transformation screening (Chapter 2).

Chapter 1

Global shifts in gene expression profiles accompanied with environmental changes in cnidarian-dinoflagellate endosymbiosis

Publication

A version of this chapter has been published as "Ishii, Y., S. Maruyama, H. Takahashi, Y. Aihara, T. Yamaguchi, K. Yamaguchi, S. Shigenobu, M. Kawata, N. Ueno, J. Minagawa. 2019. Global shifts in gene expression profiles accompanied with environmental changes in cnidarian-dinoflagellate endosymbiosis. G3: Genes, Genomes, Genetics, Early online May 16, 2019; https://doi.org/10.1534/g3.118.201012."

Abstract

Stable endosymbiotic relationships between cnidarian animals and dinoflagellate algae are vital for sustaining coral reef ecosystems. Recent studies have shown that elevated seawater temperatures can cause the collapse of their endosymbiosis, known as 'bleaching', and result in mass mortality. However, the molecular interplay between temperature responses and symbiotic states still remains unclear. To identify candidate genes relevant to the symbiotic stability, we performed transcriptomic analyses under multiple conditions using the symbiotic and apo-symbiotic (symbiont free) *Exaiptasia diaphana*, an emerging model sea anemone. Gene expression patterns showed that large parts of differentially expressed genes in response to heat stress were specific to the symbiotic state, suggesting that the host sea anemone could react to environmental changes in a symbiotic state-dependent manner. Comparative analysis of expression profiles under multiple conditions highlighted candidate genes potentially important in the symbiotic state transition under heat-induced bleaching. Many of these genes were functionally associated with carbohydrate and protein metabolisms in lysosomes. Symbiont algal genes differentially expressed *in hospite* encode proteins related to heat shock response, calcium signaling, organellar protein transport, and sugar metabolism. Our data suggest that heat stress alters gene expression in both the hosts and symbionts. In particular, heat stress may affect the lysosome-mediated degradation and transportation of substrates such as carbohydrates through the symbiosome (phagosome-derived organelle harboring symbiont) membrane, which potentially might attenuate the stability of symbiosis and lead to bleaching-associated symbiotic state transition.

1-1. Introduction

Coral reefs provide habitats for diverse marine animals, especially in oligotrophic tropical and subtropical oceans. These ecosystems rely on the stable endosymbiosis (intracellular symbiosis) between the dinoflagellate algae in the family Symbiodiniaceae and their host cnidarian animals (e.g. coral, sea anemone, and jellyfish). The symbiosis between animal hosts and Symbiodiniaceae algae likely evolved multiple times independently in different lineages (Mies *et al.* 2017). It has been proposed that recent global climate change can cause elevated water temperature and consequently the collapse of the symbiosis, known as 'bleaching' (Hughes *et al.* 2017).

Bleaching is a consequence of physiological responses to environmental changes. A major form of bleaching is induced by elevated temperature, or heat stimulus, called heat-induced bleaching (HIB); however, bleaching can also be induced by other factors (e.g. aberrant light condition, salinity, and nutrients) (Weis 2008). Two kinds of mechanisms have been proposed to be involved in the bleaching processes. One mechanism is the loss of pigments by symbionts, which causes the apparent whitening of the cnidarian host's body color but does not necessarily affect the number of symbiont cells within the host (Takahashi *et al.* 2008). The other type of bleaching is the loss of algae by the host, which can not only change the coloration but, more importantly, affect the physiological conditions of the host cells (Weis 2008).

In the cnidarian-algal relationship, symbionts are maintained in a host-derived phagosomal compartment within gastrodermal cells of the host, called symbiosome; symbiosomes have a low internal pH and are acidified by proton ATPases, as in other host cell compartments (e.g. lysosomes, endosomes, and vacuoles) (Wakefiel and Kempf 2001; Barott *et al.* 2015). Multiple routes to 'loss of algae' have been proposed, for example symbiont cell degradation within the symbiosome via fusion with lysosome and/or autophagosome, exocytosis from the host cell, and host cell death and detachment from the tissue (Weis 2008; Bieri *et al.* 2016). Although it is critical to understand how the symbiosomes are regulated under normal and stress conditions, host-symbiont communication through the symbiosome membrane at the molecular and genetic levels remains uncharacterized (Davy *et al.* 2012; Bieri *et al.* 2016).

Recently, a number of whole genome-level analyses (e.g. genomics, transcriptomics, and metabolomics) have been employed to clarify the molecular mechanisms of the symbiosis between cnidarian hosts and Symbiodiniaceae symbionts by using corals and sea anemones. Among these, the sea anemone *Exaiptasia diaphana* (formerly *Aiptasia* sp.) (Daly and Fautin 2018) is an emerging model cnidarian animal; it is experimentally feasible to induce symbiotic and apo-symbiotic (i.e. symbiont free) states reversibly by inoculation with free-living Symbiodiniaceae algae and removal of the symbionts under aberrant temperature conditions in the laboratory (Sunagawa *et al.* 2009; Lehnert *et al.* 2012; Grajales and Rodríguez 2014; Baumgarten *et al.* 2015; Weis *et al.* 2018).

Previous transcriptomic and proteomic studies reported that in host sea anemones many metabolic pathways

and cellular processes were affected by symbiosis based on the comparison between the symbiotic and apo-symbiotic states (Kuo *et al.* 2004; Rodriguez-Lanetty *et al.* 2006; Ganot *et al.* 2011; Lehnert *et al.* 2014; Oakley *et al.* 2016; Matthews *et al.* 2017). Meanwhile, a proteomic analysis using *E. diaphana* has identified a number of metabolic pathways located in cellular compartments (e.g. endoplasmic reticulum) as essential in the cellular response to heat stress (Oakley *et al.* 2017). Nevertheless, the interplay between symbiotic states and heat stress responses has not been well studied at the molecular level via genome-wide analyses.

Here, we have conducted transcriptomic analyses under different temperature conditions using *E. diaphana* in its different symbiotic states, allowing us to (1) characterize the differences in the host gene expression profiles between symbiotic and apo-symbiotic individuals in response to heat stress, (2) identify host genes and their functions, which are potentially associated with the process of heat-induced bleaching, and (3) identify symbiont gene expression changes induced by heat stress in hospite (in the host body).

1-2. Materials and methods

Strains and culture conditions

All anemones used for this study were from a clonal sea anemone *E. diaphana* strain H2 harboring a homogenous population of *Breviolum* (formerly Symbiodinium clade B) symbionts (Xiang *et al.* 2013). Symbiotic and apo-symbiotic *E. diaphana*, generous gifts from Profs. John R. Pringle and Arthur R. Grossman, were maintained at a density of 20 to 50 animals per plastic cage (12 × 22 × 13 cm, length × width × height) and fed with Artemia sp. (A&A Marine, Utah, USA) every three or four days. The symbiotic cultures were grown in circulating artificial sea water (ASW) at 25°C with fluorescent light irradiation at 60 μmol s-1 m-2 with 12 h light:12 h dark cycle. The apo-symbiotic cultures were grown in circulating ASW at 25°C without fluorescent light irradiation.

Experimental design

Incubating conditions were designed to detect changes in an early phase of heat stress responses, based on previous studies that suggested 24 h incubation at elevated temperature induced a change in the photosynthetic activity but not the number of symbiont cells in the symbiosis between *E. diaphana* and Symbiodiniaceae symbionts (Hillyer *et al.* 2016; Oakley *et al.* 2017). Prior to heat stress experiments, apo-symbiotic individuals were also incubated under the light-dark cycle for one week, and used for experiments only when they showed no sign of possible repopulation by symbiotic algae, i.e. neither algal cell bodies nor particles displaying chlorophyll fluorescence was observed under bright-field and fluorescence microscopes. During the experimental trial, symbiotic and apo-symbiotic *E. diaphana* individuals were cultured separately

at two temperatures (25°C or 33°C) for 24 h (6 h light: 12 h dark: 6 h light) at a density of 4 to 5 animals per round plastic case (4 × 9 cm, height × diameter), filled with ASW, and three individuals per treatment were sampled from a plastic case for RNAseq. To minimize contamination from Artemia RNA, anemones were not fed for 1 week prior to sampling.

Protein quantification and cell count

To measure the number of symbiont cells per host protein, each symbiotic *E. diaphana* was put into 50 µl PBS in a microtube and ground with a Biomasher II pestle (Nippi, Japan) until debris was no longer observed. Symbiont cell densities in the host cell suspension were quantified using improved Neubauer hemocytometer (Fukaekasei, Japan), with a minimum of four replicate cell counts per sample. Cell density was normalized to soluble protein content, which was assessed by using TAKARA BCA Protein Assay kit (Takara Bio, Japan) with the supernatant centrifuged (16,000 × g for 1 min) host fractions (quintuplicate measurements). Mean estimates with standard error (SEM) were calculated based on single measurements using five individuals per temperature treatment, which were separate from the samples used for transcriptome analysis.

Photosynthesis and respiration activity assay

Maximum quantum yield of photosystem II (Fv/Fm) was measured with a diving pulse amplitude modulated (PAM2500) fluorometer (Walz, Effeltrich, Germany), following a 30 min dark adaptation. PAM settings were adjusted with Ft ≤ 0.3. Photosynthesis and respiration rates were measured with a Clark-type oxygen electrode (Hansatech Instruments, Norfolk, UK) in a closed cuvette in the light at 1,000 µmol m-2 s-1 photons at 25°C. Individuals were preincubated in the dark for 10 min and then exposed to saturating light for 20 min. The respiration rate was calculated from the dark-phase oxygen consumption rate. The photosynthesis rate was calculated by subtracting the respiration rate from the light-phase oxygen evolution rate. Mean estimates with SEM were calculated based on single measurements using four individuals per temperature treatment, separate from the transcriptome analysis.

RNA extraction and sequencing

Three symbiotic or apo-symbiotic individuals per temperature condition were put into RNAlater RNA Stabilization Solution (Thermo Fisher Scientific, Massachusetts, USA) in a microtube (one individual per tube) and stored at 4°C, then the solution was replaced with 480 µl of Trizol reagent (Thermo Fisher Scientific, Massachusetts, USA). The samples were ground with a motor-assisted pestle, Biomasher II (Nippi, Japan) until debris was no longer observed. RNA extraction with Trizol reagent was conducted according to the manufacturer's instruction. The quality and quantity of RNA were verified

using Agilent RNA 6000 Nano Kit on Agilent Bioanalyzer (Agilent Technologies, California, USA) and Nanodrop spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA), respectively. One μg of total RNA from each individual was subjected to library preparation with no size selection using TruSeq RNA Sample Prep v2 Kits (Illumina, California, USA) according to the manufacturer's protocol (#15026495 Rev. D). These mRNA libraries were sequenced on Illumina Hiseq1500 with 100-mer paired-end sequence.

Transcriptome analysis

Reads from a total of 12 libraries each were obtained, trimmed, and filtered by trimmomatic option of Trinity program (Grabherr *et al.* 2011); 'PwU' output reads were used for analysis. Sequence reads were mapped against genome assemblies of *E. diaphana* (*Aiptasia* sp.) (Baumgarten et al. 2015) and *Breviolum* (formerly Symbiodinium) *minutum* (Shoguchi et al. 2013) using TopHat2 with default setting (Kim *et al.* 2013). Mapped transcripts, or fragments per kilobase of exon per million mapped fragments (FPKM) values (Mortazavi *et al.* 2008), were collected using Cufflinks Ver. 2.2.1 (Trapnell *et al.* 2010) and converted to read counts per gene using HTSeq (Anders *et al.* 2015). To obtain a visual overview of the effect of temperature treatments and host symbiotic states (symbiotic or apo-symbiotic) on global gene expression patterns, principal component analysis (PCA) was performed on the calculated FPKM values using the function "prcomp" in R (http://www/R-project.org/). Genes expressed in at least one individual were included in the PCA analysis. The count data were normalized with the R package "TCC" (Sun *et al.* 2013) and differential gene expression analysis was conducted with the "edgeR" (Robinson *et al.* 2010) analysis, implemented within the TCC package. To define differentially expressed genes (DEGs), or genes with statistically significant differences in expression between the two temperatures treatments or two host symbiotic states, the false discovery rate (FDR), or q-value, of 0.001 was used as cutoff.

Gene ontology (GO) term enrichment analysis was performed using "GOseq" package in R (Young *et al.* 2010). To annotate each *E. diaphana* gene with GO terms, BLASTp search was performed (E value cutoff, 10-4) against the Ciona intestinalis protein dataset using all of the *E. diaphana* protein dataset as query, resulting in 15279 orthologs. For symbiont genes, BLASTp search (E value cutoff, 10-4) against Arabidopsis thaliana using all of the B. minutum protein dataset as query, resulting in 15407 orthologs. Among many reference genomes available from related taxa, advantages to use the C. intestinalis and A. thaliana genomes are: (1) These species have long histories of in vivo gene function analyses and more empirical GO annotation data, (2) C. intestinalis is a relatively closely related lineage to cnidarians and useful in similarity-based homolog searches, (3) The A. thaliana genome is one of the most useful and well-documented among photosynthetic species to analyze photosynthesis-related functions. Overrepresented p-values produced by GOseq were adjusted using the Benjamini-Hochberg correction (Benjamini and Hochberg 1995). The adjusted p-value (q-value) of

0.05 was used to define enriched GO terms. Presence–absence matrix of genes associated with enriched GO terms, with dendrogram showing heatmap clustering and a table showing log fold-change (logFC) values output by TCC, was generated using "Heatmap3" package in R.

Phylogenetic analysis and localization prediction

Filtered reads of 12 libraries were de novo assembled using Trinity program (Grabherr *et al.* 2011). A contig containing 28S large subunit ribosomal RNA gene (LSU rDNA) sequence was searched by BLASTn and aligned with other sequences using the Symbiodiniaceae LSU rDNA data in a previous study (LaJeunesse *et al.* 2018). The manually curated alignment was used to reconstruct phylogenetic trees using IQ-TREE with the TIM3+F+I+G4 model which ModelFinder selected as the best model by likelihood comparison based on the Bayesian information criterion (Nguyen *et al.* 2015; Kalyaanamoorthy *et al.* 2017). The NPC2 gene sequences were translated into proteins and used for multiple sequence alignment and phylogenetic analysis as previously described (Maruyama *et al.* 2011), with the following modifications: homologous sequences automatically collected from the GenBank database were manually curated and selected for multiple alignments, and IQ-TREE was used to reconstruct phylogenetic trees using the LG+F+G4 model selected as described earlier. DHE tree was generated in the same way except using LG+I+G4. For predicting protein subcellular localization, iPSORT (Bannai *et al.* 2002) was used to predict a signal peptide or mitochondrial targeting peptide in a protein sequence, and MemPype (Pierleoni *et al.* 2011) was used to annotate eukaryotic membrane proteins.

Data availability

The dataset supporting the results of this article is included within the article and its supplemental material files. The raw data from the symbiotic and apo-symbiotic *E. diaphana* RNAseq have been submitted to the DDBJ/EMBL-EBI/GenBank under the BioProject accession number PRJDB7145.

1-3. Results

Global gene expression patterns

RNAseq produced 507,857,846 reads from apo-symbiotic ('Apo') and symbiotic ('Sym') *E. diaphana* individuals with triplicates for each culture condition (i.e. 3 Apo and 3 Sym samples under normal temperature [25°C, called 'Norm'] and elevated temperature [33° C, called 'Heat'] condition). The total number of mapped reads of the triplicates onto the dataset generated by combining all the scaffolds of the host and symbiont reference genomes were 14,573,222 reads for Sym-Norm, 13,304,783 for Sym-Heat, 20,692,119 for Apo-Norm, and 19,442,072 for Apo-Heat. Reads from co-cultured, or

'contaminating', microbes included in the raw data were filtered out in this mapping process. Overall, we obtained the FPKM and count values of genes in the *E. diaphana* genomes under each of four conditions (i.e. Apo-Norm, Apo-Heat, Sym-Norm, and Sym-Heat) and the B. minutum genome for two conditions (i.e. Norm and Heat). In the symbiotic and apo-symbiotic samples, the proportions of reads mapped onto the symbiont genome sequences were about 10% and less than 1%, respectively.

We conducted principal component analysis (PCA) of the gene expression patterns using the FPKM data mapped to the host *E. diaphana* or the B. minutum genome. In *E. diaphana*, the first principle component (PC1) represented the effect of symbiotic states (i.e. apo-symbiotic or symbiotic), whereas PC2 represented those of temperature on gene expression levels (Figure 1-1A). The analysis separated the four conditions with no overlap (Figure 1-1A). In symbionts, the gene expression patterns in Norm and Heat thermal treatments were weakly separated along the PC2 axis (Figure 1-1B).

For the host sea anemone transcriptome analysis, the *E. diaphana* genome was used as a reference (Baumgarten *et al.* 2015). For the symbiont transcriptome, the *B. minutum* (formerly *S. minutum*, clade B) genome (Shoguchi *et al.* 2013) was used as a reference, as we found only one contig matching to 28S large subunit ribosomal RNA gene in our data, which formed a clade with *Breviolum* sequences (Figure S1-1).

The exposure to the elevated temperature for one day resulted in no apparent symptom of bleaching and no significant decline in the number of algal cells (Figure S1-2A), maximum quantum yield of photosystem II (Figure S1-2B), or photosynthesis rates (Figure S1-2C); however, respiration rates were decreased after the heat treatment (Figure S1-2D).

E. diaphana DEGs

To detect genes differentially expressed between the test conditions in reference to previous studies, we used the FDR of 0.001 as the threshold, which is more stringent than the values used in other studies (e.g. FDR of 0.05). For reference purpose, we compared the expression patterns of the Npc2-type sterol transporter gene family. To our knowledge, NPC2 is one of the few examples of which the gene expression patterns were differentially regulated dependent on the symbiosis state in multiple cnidarian species in previous studies from multiple research groups (Lehnert *et al.* 2014; Dani *et al.* 2014; Baumgarten *et al.* 2015); and references therein). In our data, out of six NPC2 homologs, five genes were differentially expressed and all up-regulated in the Sym-Norm relative to the Apo-Norm condition (Figure S1-3, S1-4); thereby, results were consistent with previous reports (Lehnert *et al.* 2014). Furthermore, four genes among those five were significantly down-regulated in the Sym-Heat compared to the Sym-Norm group (Figure S1-3, S1-4).

By comparing Apo-Norm and Apo-Heat test groups, we identified 594 genes as DEGs, whereas the Sym-

Norm vs Sym-Heat comparison resulted in 927 DEGs (Figure 1-2), which we call here two groups of heat-responsive DEGs (HR-DEGs). Between both groups, 190 genes were shared, called 'shared HR-DEGs'. Gene expression regulation of the shared HR-DEGs were conserved between the symbiotic states, as the majority of HR-DEGs were down-regulated at an elevated temperature compared to the normal temperature in symbiotic and apo-symbiotic individuals (Figure S1-5). Heat shock proteins (HSPs) have been reported to be activated in some coral-symbiont systems under elevated temperature (Desalvo *et al.* 2008; 2010). In *E. diaphana*, only H90A1 was a shared HR-DEG, while HS71A, HSP7C, HSP97, and HSP7C were detected as unique HR-DEGs only found in the symbiotic individuals, and CH10 and AHSA1 only found in the apo-symbiotic state (Figure S1-6).

GO term enrichment analysis detected eight GO terms enriched in 190 shared HR-DEGs and 13 terms in 737 HR-DEGs unique to the symbiotic individuals (Figure 1-2, Table S1-1, Table S1-2). No GO term was enriched in 404 HR-DEGs unique to the apo-symbiotic individuals (Figure 1-2, Table S1-3). We analyzed these genes associated with the enriched GO terms by mapping them onto presence–absence matrices. The enriched GO terms of 190 shared HR-DEGs could be divided into four groups by heat map clustering for descriptive purposes (Figure S1-5). Groups 1, 2, 3 and 4 were related to methylation, protein folding in ER, transmembrane transport, and oxidation-reduction, respectively. The enriched GO terms of the HR-DEGs unique to the symbiotic individuals were partially overlapped with the ones for the shared HR-DEGs, i.e. oxidation-reduction (Group 1), transmembrane transport (Group 5) (Figure S1-7).

Considering HIB can have a major impact on coral bleaching in nature (Weis *et al.* 2018), we collected the DEGs that could potentially be associated with the HIB process. The four culture conditions introduced in this study can be assumed to mimic steps of a HIB process (Fig. 1-3A): Sym-Norm (steady state prior to HIB), Sym-Heat (temperature elevation), Apo-Heat (heat-induced collapse of symbiosis), and Apo-Norm (steady state after HIB). In addition, the expressions of genes relevant to HIB can be considered to be altered irreversibly as the process progresses. We detected DEGs of which the expression levels were changed in the same direction (i.e. either up- or down-regulated) in response to temperature elevation (Sym-Heat relative to Sym-Norm), symbiotic state transition (Apo-Norm relative to Sym-Norm), and a combination of those (Apo-Heat relative to Sym-Norm); these were called HIB-associated (HIBA) genes (Figure 1-3A). We identified 292 HIBA genes (Figure 1-3B, Table S1-4) and detected nine enriched GO terms associated with the HIBA genes, which were classified into four groups: transporter (Group 1), oxidation-reduction (Group 2), lysosome (Group 3), and carbohydrate metabolism (Group 4) (Figure 1-3C).

Symbiont DEGs

By comparing the *Breviolum* spp. symbiont Norm and Heat samples, we identified 124 genes as HR-DEGs in the

symbionts and conducted the GO term enrichment analysis of these genes based on the A. thaliana genome annotation data (Table S1-5), resulting in 12 terms detected. Although a number of groupings were recognized on the presence–absence matrix (Figure 1-4), many of the symbiont HR-DEGs were associated with multiple enriched GO terms and the classifications were not straightforward. Group 1 was associated with response to cadmium ion, while Group 2, 3, 4, and 5 were heat response, cytosol, ATP-binding, and stress response, respectively. A number of the symbiont HR-DEGs were not associated with any enriched GO terms, but potentially relevant to heat stress response in the symbionts (See Discussion) (Table S1-5).

1-4. Discussion

The host transcriptomic differences between symbiotic and apo-symbiotic states in response to heat stress

Our results show that heat stress responses, at least at the transcriptomic level, are different depending on the symbiotic states. Incubation under the elevated temperature conditions (33°C for 24 h) led to no apparent indication of loss of algae (Figure S1-2). Considering previous studies showing that the number of symbiont cells decreased over time with incubation at elevated temperatures (Dunn *et al.* 2004; Hawkins *et al.* 2013), the Sym-Heat transcriptome in this study is most likely to reflect the very early phase of environmental response by the host *E. diaphana* prior to, rather than in process of, bleaching. Although the *E. diaphana* individuals used for RNAseq were cultured by treatment and thus the conditions were not ideally randomized, *HSP* gene expression patterns (Figure S1-6), which were known to be heat-responsive, suggest that the effect of elevated temperature was detected in this analysis. Although the possibility that the effect of container (e.g. water amount, specific density) affected the gene expression could not be denied, we postulated that it was limited. The transcriptome profiling analyses revealed that both the symbiotic states and the thermal treatment had clear and substantially divergent effects on global gene expression patterns in the host (Figure 1-1A, 1-2), which is consistent with previous reports that symbiotic states are associated with different transcriptomic, proteomic, and metabolic profiles under normal growth conditions (Mitchelmore *et al.* 2003; Lehnert *et al.* 2014; Oakley *et al.* 2016).

The differential expression patterns of NPC2-type sterol transporter genes further corroborated that our experimental settings were comparable to previous studies, which showed that a gene family member *NPC2D* was up-regulated in the symbiotic state compared to the apo-symbiotic state in *E. diaphana* (Lehnert *et al.* 2014; Baumgarten *et al.* 2015) (Figure S1-3). In addition, *E. diaphana NPC2B, C, D,* and *F* were closely related to *NPC2-d* in the snakelocks anemone *Anemonia viridis* (Figure S1-4), which was shown to be preferentially accumulated in the gastrodermal cells at the mRNA and protein levels and significantly down-regulated in response to heat stress (Dani *et al.* 2014). Most *NPC2* genes were differentially expressed, except between Apo-Norm and Apo-Heat (Figure S1-3, S1-4). This may be related to

the finding that the *NPC2* gene expression was less sensitive to heat in the symbiont-free epidermis cells in *A. viridis* (Dani *et al.* 2014).

Previous studies showed that well-known stress response genes encoding HSPs were differentially expressed in response to heat stress in the corals *Orbicella* (formerly *Montastraea*) *faveolata* (Desalvo *et al.* 2008) and *Acropora palmata* (Desalvo *et al.* 2010), but no significant DEGs were detected in the temperate sea anemone *Anthopleura elegantissima* (Richier *et al.* 2008). In the case of *E. diaphana*, a major heat stress responsive chaperone *H90A1* (Picard 2002) was up-regulated regardless of symbiotic state, while different HSP genes were differentially expressed solely in either symbiotic state (Figure S1-6). Overall, these results imply the presence of multiple pathways for regulating the expression of genes, including HR-DEGs, in a symbiotic state- and/or species-dependent manner.

To further investigate the gene expression regulation and functional properties of HR-DEGs, we conducted GO term enrichment analysis. (Figure 1-2, Figure S1-5, Figure S1-7). The results suggest that some functions (e.g. oxidation-reduction process and transmembrane transporter activity) are enriched in both the shared HR-DEGs and symbiotic-specific HR-DEGs. For instance, as previous studies showed that sea anemone and coral expel living symbionts as a pellet wrapped with the host mucus (Steele 1977) by exocytosis (Davy *et al.* 2012), heat stress may undermine the regulation of mucus rearrangement and resynthesis using genes related to extracellular matrix functioning (e.g. collagen alpha *COLA1* and *CO4A2*, Fibrillin *FBN1* and *FBN2*) when the host accommodates symbionts (Figure 1-2, Figure S1-7).

Genes and functions associated with 'heat-induced bleaching'

GO term enrichment analysis using the HIBA genes identified four functional groupings (Figure 1-3C). Considering that the HIBA genes were important candidates for their role in bleaching, these results could provide insights into the cellular and molecular functions involved in this process. Most of the genes linked with carbohydrate metabolism (e.g. Alpha-N-acetylgalactosaminidase *NAGAB*, Lysosomal alpha-mannosidase *MA2B1*) were involved in degradation and/or modification of complex carbohydrates such as *N*-linked glycosylation of glycoprotein, which are generally generated in the Golgi apparatus and transported to a lysosome (Table S1-4) (Winchester 2005). Meanwhile, many of the genes associated with the term 'lysosome' were related to lysosomal protein modification or protease activities, e.g. Beta-glucuronidase *BGLR,* Palmitoyl-protein thioesterase 1 *PPT1*, Cathepsin proteases *CATB, CATL, CYSP* (Table S1-4). A lysosome is a versatile organelle and in normal conditions it is likely that the HIBA genes are involved in multiple functions such as regulating turnover rates of host proteins. In the HIB process, lysosomal degradation and modification functions may be suppressed by down-regulating the HIBA genes (Figure 1-5).

In the cnidarian-algal symbiosis, symbionts are maintained in symbiosomes within gastrodermal cells of the

host (Wakefiel and Kempf 2001). Symbiodiniaceae symbionts in sea anemones and corals can be digested in certain conditions (Bieri *et al.* 2016), presumably via fusion with the lysosome- and autophagy-mediated degradation (Weis 2008). Additionally, it was reported that in the *Hydra-Chlorella* symbiosis, unhealthy symbionts might be swept out via fusion of lysosomes with symbiosomes (Hohman *et al.* 1982). In regulating lysosome-phagosome fusion, the Rab GTPase gene family, an important regulator of vesicular trafficking (Schwartz *et al.* 2007), has been proposed to play key roles. Previous studies suggested that Rab family proteins were localized in phagosomes and are possibly involved in the exclusion and maintenance of symbionts in *Aiptasia pulchella* (a synonym of *E. diaphana*) (Chen *et al.* 2005; Hong *et al.* 2009). In our data, the HIBA genes included a gene encoding Rab32, which is a regulator of the lysosomal enzyme recruitment to phagosome (Seto *et al.* 2011); the transcription regulation may be relevant to the symbiosome maintenance (Figure S1-8).

Another functional group in the HIBA genes is 'transporter,' including facilitated glucose transporters *GTR1* (*GLUT1*) and *GTR8* (*GLUT8*), monocarboxylate transporters *MOT8* and *MOT10. GTR8* is proposed to be a potential symbiosome-localized transporter that transfers glucose from the symbiont to the host (Sproles *et al.* 2018), and perhaps other HIBA transporters may also be associated with phagosome and/or symbiosome. Proton oligopeptide cotransporter SLC15A4 and aquaporin-like MIP proteins were predicted to be localized to internal membrane by MemPype, while seven other transporter candidates were not. Further biochemical studies are needed.

The other HIBA group 'oxidation-reduction' included two closely related *DHE3* genes encoding a cnidarian-specific subtype of glutamate dehydrogenase (Figure S1-9A), which are only distantly related to another homolog (AIPGENE3776) belonging to the canonical mitochondrion-targeted *DHE3* gene family widely conserved in eukaryotes. Notably, the transcription of these two genes was regulated in the opposite direction (Figure 1-3C), and iPSORT program predicted a mitochondrial targeting signal in the N-terminal amino acid sequence of up-regulated gene, but not in the down-regulated one (Figure S1-9B), suggesting differential transcriptional regulation for closely related homologs localized in different intracellular compartments (Figure 1-5) (Mastorodemos *et al.* 2009).

The symbiont transcriptomic responses induced by heat stress *in hospite*

Our results showed that some *HSP* gene family members (e.g. *HSP70-14, HSP90-1, 90-2, 90-4*) constitute major components of the symbiont HR-DEGs when *Breviolum* symbionts were hosted by *E. diaphana* (Figure 1-4, Table S1-5). A previous study demonstrated that thermal stress did not induce substantial global changes in the transcriptomes of the symbiont *Durusdinium* spp. (clade D *Symbiodinium*) colonized in the coral *Acropora hyacinthus*; whereas, a few genes encoding HSPs were weakly differentially expressed in response to heat stress (Barshis *et al.* 2014). Another study showed that *HSP70* and *HSP90* genes in *Cladocopium* (clade C) colonized in the coral *Acropora millepora* were differentially

expressed in response to heat stress, depending on the acuteness of the heat treatment (Rosic *et al.* 2011). These results collectively suggest that symbiont *HSP* gene expression was differentially regulated by multiple factors, including environmental and phylogenetic constraints. *FKB62,* another symbiont HR-DEG, encoded a peptidyl prolyl isomerases FKBP62 (Figure 1-4, Table S1-5), which plays a role in high temperature tolerance by interacting with HSP90.1 and stabilizing small HSPs in *Arabidopsis* (Meiri and Breiman 2009). Our data showed that *FKB62* as well as many *HSPs* in the symbionts were down-regulated in response to heat (Figure 1-4), in contrast to the up-regulation of host *HSPs* (Figure S1-6), implying that symbiont *HSPs* might negatively regulate the heat responsive gene expression, as proposed in a previous study (Rosic *et al.* 2011).

Found in the symbiont HR-DEG were a component of the inner chloroplast membrane translocon (TIC) complex *TIC20*, chloroplastic/mitochondrial presequence protease 1 *PREP1*, chloroplastic chaperone proteins *ClpC1* and *ClpB3,* a mitochondrial Lon protease homolog *LON1*, and an import receptor for peroxisomal-targeting signal peptide *PEX5*; this result suggests that proper regulation of protein import from cytosol to organellar compartments was inhibited in symbionts under heat stress. In cytosol, the following symbiont HR-DEGs may be directly or indirectly involved in calcium signaling and affected by heat stress: glutamate-gated cation channel proteins *GLR3.5* and *GLR3.7*, calcium-dependent protein kinases *CPK19* and *CPK23*, a plasma membrane-localized ammonium transporter *AMT1-3*, a carbamoyl phosphate synthetase B *CARB* (Table S1-5) (Michaeli and Fromm 2015). As ammonium assimilation is a core process in the nitrogen cycling and amino acid relocation between cnidarian hosts and their symbionts (Pernice *et al.* 2012) , genes involved in this process such as *AMT1-3* and *CARB* may also play a key role in molecular interactions in the nitrogen-limited oligotrophic oceans, via balancing the carbon/nitrogen ratio in the symbiont cells (Houlbrèque and Pagès 2009) .

The symbiont HR-DEGs contain two sugar metabolism-related genes. One is a gene encoding a nucleotide-sugar transporter *GONST3*, which may function in the import of nucleotide-sugar from cytosol to the Golgi apparatus for downstream glycosylation reactions. The other encodes a sucrose-phosphate synthase family protein *SPS4*, which might be related to photosynthetic sucrose synthesis. In the cnidarian-algal symbiosis, it is suggested that sugar, more specifically glucose, is an important component for not only the supply of photosynthesized carbohydrates from symbiont to host (Burriesci *et al.* 2012) but also for the recognition of symbionts by the host (Takeuchi *et al.* 2017; Huang *et al.* 2017). Our results raise a possibility that cytosolic sugar metabolism and Golgi apparatus-mediated glycosylation of proteins and/or cell wall components may be susceptible to stress and damage when symbionts are exposed to heat *in hospite* (Figure 1-5).

A hypothesis on the molecular interplay between host and symbiont under heat stress

Our data pinpoint that, in addition to the differences of steady state transcriptomes, cnidarian hosts possessing the same genetic background can respond to the same environmental changes, such as heat stress, in very different ways depending on their symbiotic state (Figures 1-1, 1-2). Furthermore, we identified HIBA genes associated with the symbiotic state transition and showed novel predicted functions of potential importance in symbiosome maintenance (Figure 1-5).

One plausible hypothesis is that the HIBA genes play key roles in lysosomal (or symbiosomal) degradation and modification of glycoproteins at the symbiont cell surface (Winchester 2005) and thereby affecting the symbiosis stability under heat stress (Figure 1-5). Previous studies suggested that lectin proteins capable of binding the glucose moiety might be involved in the recognition of Symbiodiniaceae symbionts by the host coral *Acropora tenuis* (Takeuchi *et al.* 2017). Furthermore, a glycoprotein was characterized as the first Symbiodiniaceae protein and was localized at the cell surface, expressed exclusively when the symbiont was colonized within the host (Huang *et al.* 2017). In the HIB process, the altered transport rate of degraded metabolites to the host cytosol may work as a negative feedback signal for the subsequent decrease of metabolite flow (Davy *et al.* 2012). Further investigation of the molecular interaction between host and symbiont, presumably mediated via glycoprotein metabolism in lysosomes and symbiosomes, will be key to understanding what signal can trigger the collapse of symbiosis.

Figures 1

Figure 1-1. Global gene expression patterns in *E. diaphana* and the symbionts

2. PCA of the *Exaiptasia diaphana* transcriptomes. B. PCA of the symbiont transcriptomes *in hospite*.

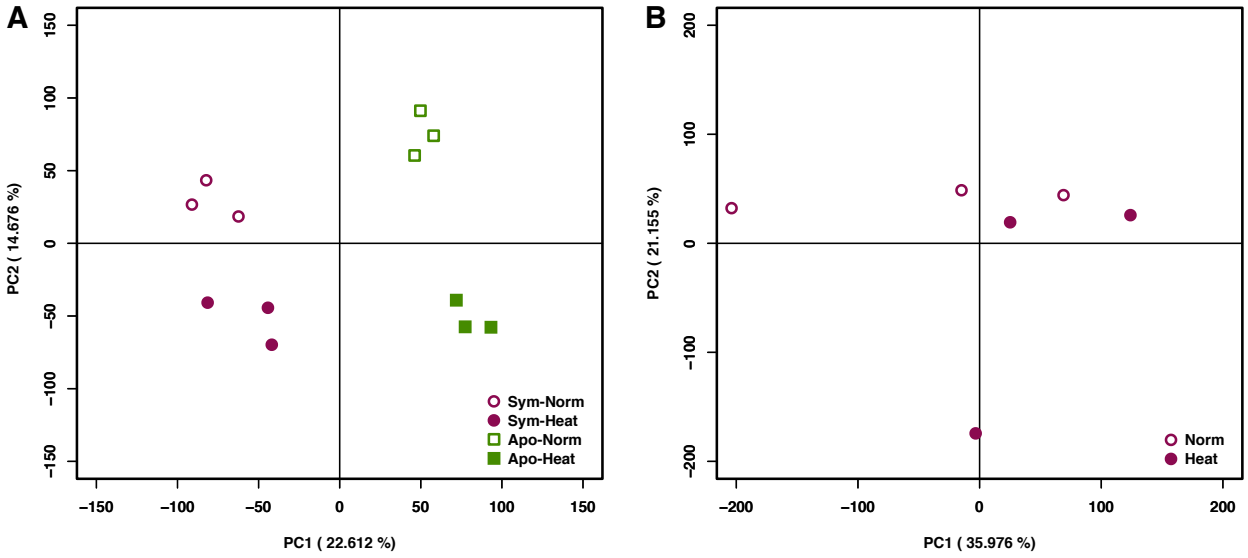


Figure 1-2. *E. diaphana* genes differentially expressed under heat stress in symbiotic and apo-symbiotic states

Venn diagram presents comparisons of the numbers of DEGs in each symbiotic state. Enriched GO terms also are shown for each compartment of the diagram. BP, biological process; CC, cellular component; MF, molecular function.

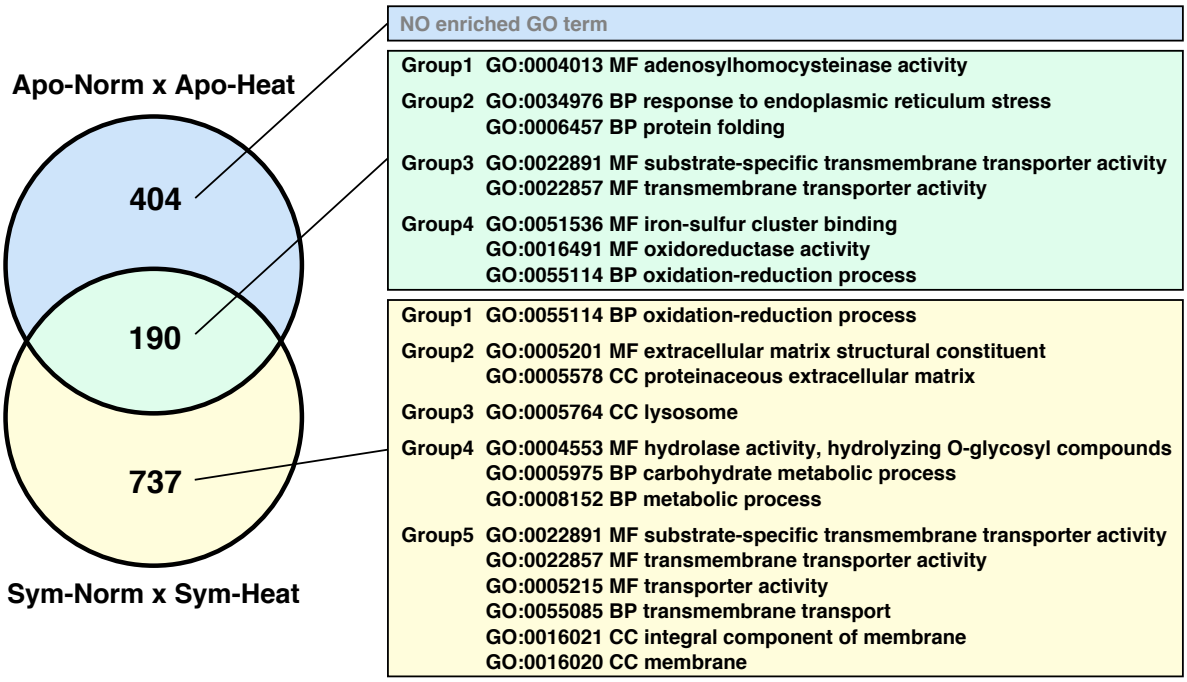


Figure 1-3. Expression patterns of HIBA genes

Conceptual representation of HIBA gene definition. HIBA genes are defined as DEGs of which all the expression levels in Sym-Heat, Apo-Heat, and Apo-Norm have been changed in the same direction in comparison with Sym-Norm. Assumed bleaching conditions are shown along with sample conditions. Asterisks indicate significantly differential gene expression. B. Venn diagram presents the numbers of DEGs in multiple comparisons. 'Up' and 'Down' DEGs indicate the ones up-regulated and down-regulated relative to Sym-Norm, respectively. C. Presence–absence matrix of HIBA genes associated with enriched GO terms. HIBA genes are shown with 'AIPGENE' gene IDs and putatively annotated gene names on the vertical axis, with a heat map showing gene expression levels as log FC values. Gene IDs shown in magenta and teal blue are up- and down-regulated genes relative to Sym-Norm. Enriched GO terms are shown with GO ID, GO category and description on the horizontal axis, with a clustering based on the genes presented in each GO term column. Closed and open cells indicate the presence and absence of the association with GO terms.

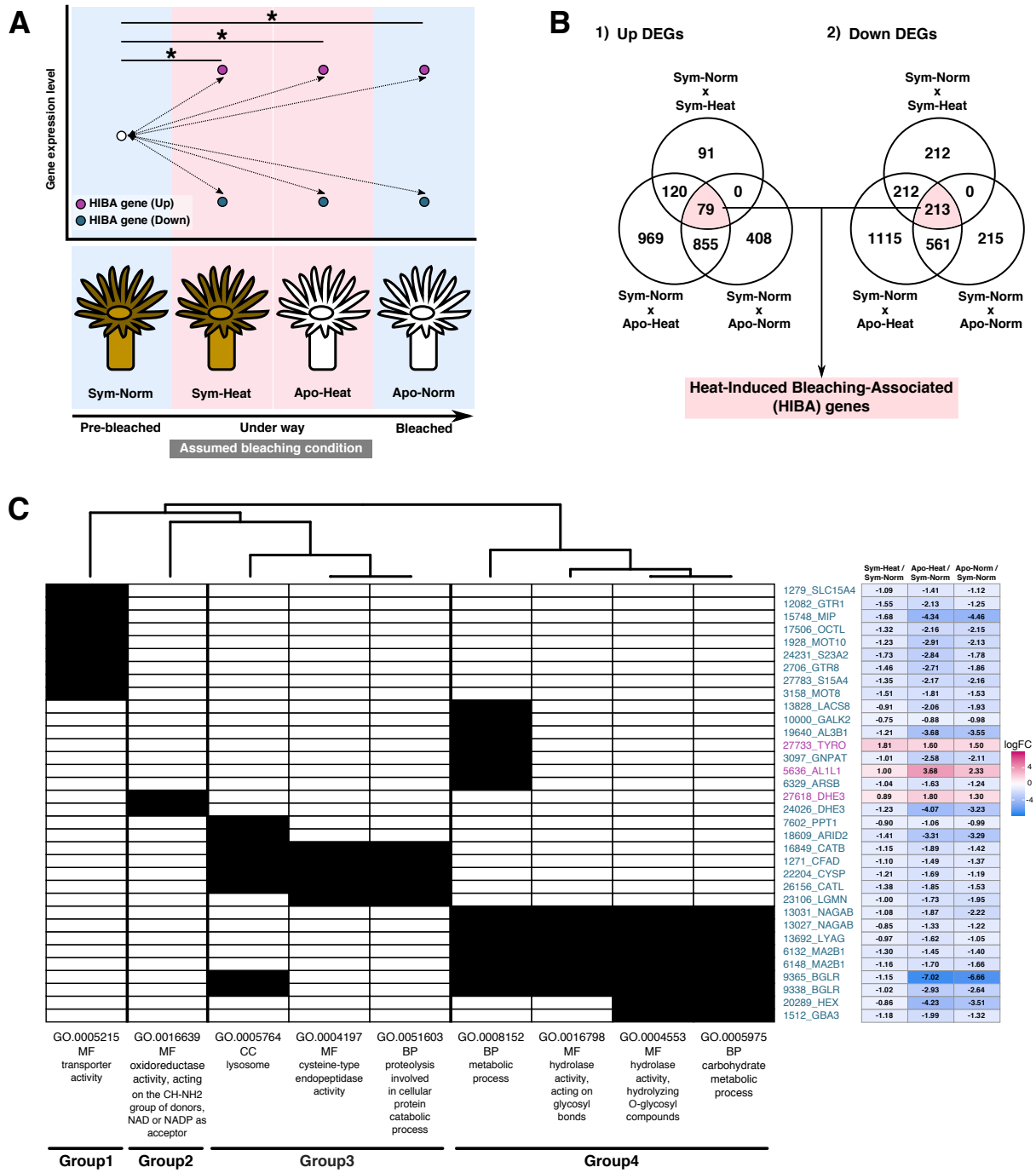

A
Gene expression level
HIBA gene (Up)
HIBA gene (Down)
Sym-Norm
Sym-Heat
Apo-Heat
Apo-Norm
Pre-bleached
Under way
Bleached
Assumed bleaching condition
B
1) Up DEGs
2) Down DEGs
Sym-Norm x Sym-Heat
91
120
0
79
969
855
408
Sym-Norm x Apo-Heat
Sym-Norm x Apo-Norm
212
212
0
213
1115
561
215
Heat-Induced Bleaching-Associated (HIBA) genes
C
Sym-Heat / Sym-Norm
Apo-Heat / Sym-Norm
Apo-Norm / Sym-Norm
1279_SLC15A4 -1.09 -1.41 -1.12
12082_GTR1 -1.55 -2.13 -1.25
15748_MIP -1.68 -4.34 -4.46
17506_OCTL -1.32 -2.16 -2.15
1928_MOT10 -1.23 -2.91 -2.13
24231_S23A2 -1.73 -2.84 -1.78
2706_GTR8 -1.46 -2.71 -1.86
27783_S15A4 -1.35 -2.17 -2.16
3158_MOT8 -1.51 -1.81 -1.53
13828_LACS8 -0.91 -2.06 -1.93
10000_GALK2 -0.75 -0.88 -0.98
19640_AL3B1 -1.21 -3.68 -3.55
27733_TYRO 1.81 1.60 1.50
3097_GNPAT -1.01 -2.58 -2.11
5636_AL1L1 1.00 3.68 2.33
6329_ARSB -1.04 -1.63 -1.24
27618_DHE3 0.89 1.80 1.30
24026_DHE3 -1.23 -4.07 -3.23
7602_PPT1 -0.90 -1.06 -0.99
18609_ARID2 -1.41 -3.31 -3.29
16849_CATB -1.15 -1.89 -1.42
1271_CFAD -1.10 -1.49 -1.37
22204_CYSP -1.21 -1.69 -1.19
26156_CATL -1.38 -1.85 -1.53
23106_LGMN -1.00 -1.73 -1.95
13031_NAGAB -1.08 -1.87 -2.22
13027_NAGAB -0.85 -1.33 -1.22
13692_LYAG -0.97 -1.62 -1.05
6132_MA2B1 -1.30 -1.45 -1.40
6148_MA2B1 -1.16 -1.70 -1.66
9365_BGLR -1.15 -7.02 -6.66
9338_BGLR -1.02 -2.93 -2.64
20289_HEX -0.86 -4.23 -3.51
1512_GBA3 -1.18 -1.99 -1.32
logFC
4
0
-4
GO.0005215 MF transporter activity
GO.0016639 MF oxidoreductase activity, acting on the CH-NH2 group of donors, NAD or NADP as acceptor
GO.0005764 CC lysosome
GO.0004197 MF cysteine-type endopeptidase activity
GO.0051603 BP proteolysis involved in cellular protein catabolic process
GO.0008152 BP metabolic process
GO.0016798 MF hydrolase activity, acting on glycosyl bonds
GO.0004553 MF hydrolase activity, hydrolyzing O-glycosyl compounds
GO.0005975 BP carbohydrate metabolic process
Group1
Group2
Group3
Group4

Figure 1-4. Presence–absence matrix and expression levels of HR-DEGs in symbionts

Symbiont HR-DEGs associated with enriched GO terms are shown as in Figure 3C, with the *B. minutum* gene IDs and putatively annotated gene names. Gene IDs shown in magenta and teal blue are up- and down-regulated genes relative to symbiont Norm, respectively.

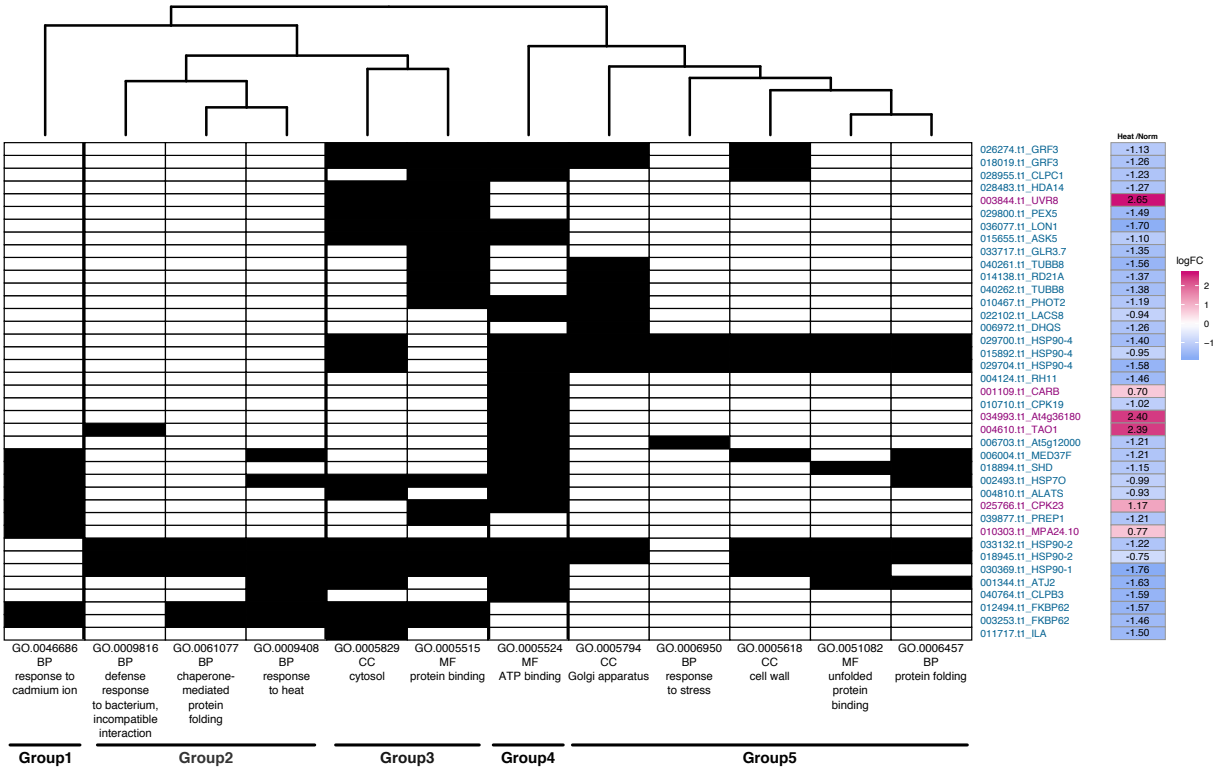
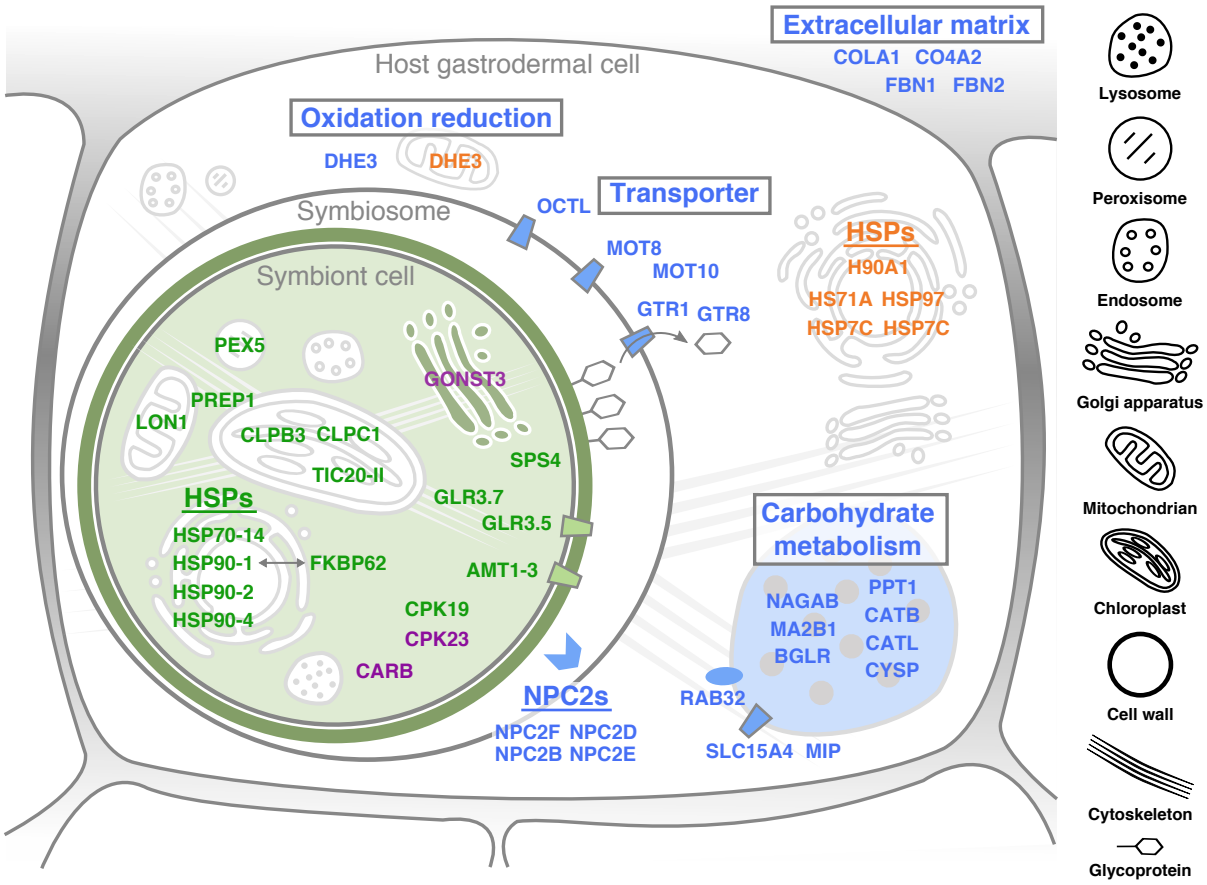


Figure 1-5. A model of the molecular interplay between host and symbiont under heat stress

Text color indicates up-regulated (orange and purple) and down-regulated (blue and green) expression in the host and symbiont cells, respectively. Field color and boxed text indicate organelles and gene functions potentially involved in heat stress response, respectively (see Discussion). Data are from Figure 1-3, 1-4, S1-5, S1-7, Table S1-1, S1-2, S1-4, S1-5.



Chapter 2

Isolation of uracil auxotroph mutants of coral symbiont alga for symbiosis studies

Publication

A version of this chapter has been published as "Ishii, Y., S. Maruyama, K. Fujimura-Kamada, N. Kutsuna, S. Takahashi, M. Kawata, J. Minagawa. 2018. Isolation of uracil auxotroph mutants of coral symbiont alga for symbiosis studies. Scientific Reports., 8: 3237."

Abstract

Coral reef ecosystems rely on stable symbiotic relationship between the dinoflagellate *Symbiodinium* spp. and host cnidarian animals. The collapse of such symbiosis could cause coral 'bleaching' and subsequent host death. Despite huge interest on *Symbiodinium*, lack of mutant strains and readily available genetic tools have hampered molecular research. A major issue was the tolerance to marker antibiotics. Here, we isolated *Symbiodinium* mutants requiring uracil for growth, and hence, useful in transformation screening. We cultured *Symbiodinium* spp. cells in the presence of 5-fluoroorotic acid (5FOA), which inhibits the growth of cells expressing *URA3* encoding orotidine-5'-monophosphate decarboxylase, and isolated cells that require uracil for growth. Sequence analyses and genetic complementation tests using yeast demonstrated that one of the mutant cell lines had a point mutation in *URA3*, resulting in a splicing error at an unusual exon–intron junction, and consequently, loss of enzyme activity. This mutant could maintain a symbiotic relationship with the model sea anemone *Exaiptasia pallida* only in sea water containing uracil. Results show that the *URA3* mutant will be a useful tool for screening *Symbiodinium* transformants, both ex and in *hospite*, as survival in the absence of uracil is possible only upon successful introduction of *URA3*.

2-1. Introduction

The dinoflagellate *Symbiodinium* spp. are known to sustain a stable symbiotic relationship with cnidarian animals (e.g. coral, sea anemone, jellyfish) by endosymbiosis in the gastroderm (endoderm) cells of host cnidarian animals (Davy *et al.* 2012). Ecologically, *Symbiodinium* is a key primary producer for sustaining the coral reef ecosystems in the oligotrophic tropical and subtropical ocean, and much of the photosynthetically fixed carbon by symbionts is provided to the host coral (Brown 1997). Collapse of the coral-algal symbiosis, which is known as 'bleaching', often leads to death of the host corals, causing destructive damage on the coral reef ecology (Brown 1997).

From a taxonomical perspective, although *Symbiodinium* spp. can be classified into a number of 'clades' by means of molecular phylogeny (Pochon and Gates 2010), all these clades lack conspicuous morphological traits applicable for species-level classification. Recently, advances in sequencing technology have revealed the diversity of *Symbiodinium* across a range of coral reefs and other marine environments. Previous studies suggested that the specificity of the *Symbiodinium*-cnidarian symbioses was dependent on the size of the algal symbiont (Biquand *et al.* 2017), and that the symbiont specificity of corals increased (i.e. fewer *Symbiodinium* types can be associated with corals) as the host coral grew (Cumbo *et al.* 2013). On the other hand, no substantial change on the symbiont specificity was observed in the model sea anemone *Exaiptasia pallida* (formerly *Aiptasia* sp.) (Hambleton *et al.* 2014).

In spite of the accumulation of genomic and multi-omics information on cnidarian-algal symbiosis (Mohamed *et al.* 2016), genetic tools for characterizing functions of genes that are involved in such symbiosis are still very limited and not readily available. Two independent studies on gene delivery into the *Symbiodinium* cells have been published. The first report by ten Lohuis and Miller discusses successful delivery of external DNA molecules using silicon carbide whiskers (Lohuis and Miller 1998). Seventeen years later, Ortiz-Matamoros and colleagues reported transient expression of exogenous genes delivered into *S. microadriaticum* subsp. *Microadriaticum* strain S. KB8, *Symbiodinium* sp. strain Mf11.5b.1, and the genome-sequenced strain *Symbiodinium kawagutii* (Lin *et al.* 2015), using polyethylene-glycol with glass beads (Ortiz-Matamoros and Villanueva 2015) or the terrestrial bacterium *Agrobacterium tumefaciens*, which has been widely used for transformation of land plants (Ortiz-Matamoros *et al.* 2015). However, further elaboration of methods for gene delivery into *Symbiodinium* cells is clearly needed: No follow-up studies have been published using the methods developed by ten Lohuis and Miller(Lohuis and Miller 1998) and, although it was shown that transient gene introduction methods used for land plants could also be applicable to *Symbiodinium*, no stable transformant lines have been reported (Ortiz-Matamoros *et al.* 2015).

Towards developing a genetic tool for the coral symbiotic dinoflagellate *Symbiodinium* sp., we conducted antibiotic screening experiments and isolated a nutrient (uracil)-requiring *Symbiodinium* mutant. We show that the cell

growth could be switched on and off by replacing the media, and that the growth switching was inducible in and ex *hospite*.

2-2. Materials and methods

Culturing and screening methods

Symbiodinium strain SSB01, a generous gift from Profs. John R. Pringle and Arthur R. Grossman, was an axenic unialgal strain closely related to the genome-sequenced strain *S. minutum* Mf1.05b (clade B) (Xiang *et al.* 2013; Shoguchi *et al.* 2013). SSB01 was maintained at 25 °C in Marine Broth (MB) medium containing 33.5 g L^{-1} of MB (Difco Laboratories, New Jersey, USA), 250 mg L^{-1} of Daigo's IMK Medium (Nihon Pharmaceutical, Japan), and PSN (Gibco, Thermo Fisher Scientific, Massachusetts, USA) where the final concentrations of penicillin, streptomycin and neomycin were 0.01, 0.01, 0.02 mg ml^{-1}, respectively. Light was provided at an irradiance of approximately 100 μmol photons m^{-2} s^{-1} in a 12 h light: 12 h dark cycle. IMK medium, which contained 33.5 g L^{-1} of Sea salt (Sigma-Aldrich, Merck Millipore, Germany), and 250 mg L^{-1} of Daigo's IMK Medium, was also used in some experiments.

To obtain spontaneous 5FOA-resistant mutants, approximately 1 to 5 × 10^7 cells of SSB01 were concentrated by centrifugation and spread on each MB plate (90 mm dish) containing 1% Agar (Wako Pure Chemical Industries, Japan) and 200 μg L^{-1} of 5FOA. After all, four independent clones were isolated from eight screening plates after 2 months of incubation under a 12 h light: 12 h dark cycle, with lighting provided at an irradiance of 5-20 μmol photons m^{-2} s^{-1}, at 25 ˚C. After subsequent culturing, we designated these clones showing stable growth in the liquid media as T01 (Tohoku University strain 01), T22, T23 and T29.

To examine the 5FOA-resistant and uracil-requiring phenotypes, we grew SSB01 and T01, T22, T23 on MB and IMK plates either containing or lacking 5FOA. The cells were spotted at the concentration of 2.5 × 10^2 cells/10 μl on MB plate followed by 4-week incubation, and 8.75 × 10^2 cells/10 μl on IMK plate followed by 10-week incubation. To compare the growth of SSB01 and T01, we cultured these strains in liquid media. For liquid cultures using MB and artificial seawater (ASW) using Instant Ocean sea salt (Aquarium systems, France), the cells were inoculated in 20 ml of media (3.125 × 10^5 cells ml^{-1}) in T25 flasks. The cell growth was measured by counting the number of cells using TC20 Automated Cell Counter (Bio-Rad Laboratories, California, USA). Student's two-tailed t-tests were performed to test whether the growth rates significantly differed between the strains and treatments.

For antibiotics resistance tests, we grew *Symbiodinium* strains CCMP 830 (clade B), 2429 (clade A), 2434 (clade B) and 2455 (clade F) on IMK plate, where the concentration of the sea salt was reduced to half to increase the sensitivity to antibiotics, containing 1% agar and 0.1 mg ml^{-1} ampicillin plus either of the following antibiotics; 0.2 mg ml^{-1} kanamycin, 0.4 mg ml^{-1} neomycin, 1.0 mg ml^{-1} streptomycin, 0.2 mg ml^{-1} paromomycin, 0.2 mg ml^{-1} nourseothricin

(Sigma-Aldrich, Merck Millipore, Germany) and 0.2 mg ml^{-1} zeocin (Thermo Fisher Scientific, Massachusetts, USA).

RNA extraction and determining the URA3 transcript sequence

RNA was extracted from SSB01 and T01, T22, T23, T29. Cells were collected by centrifugation at 13,000 × g for 1 min at 25 °C. Culture medium was discarded, and the cell pellet was re-suspended in 500 µl of TRIzol (Thermo Fisher Scientific, Massachusetts, USA) and approximately 20 µl of Glass beads (Sigma-Aldrich, Merck Millipore, Germany) were added and mixed well by vortexing, followed by addition of an equal volume of chloroform and vortexing. After centrifugation and separation, the aqueous phase was used to purify total RNA by Rneasy Mini kit (Qiagen GmBH, Germany) following the manufacturer's protocol. For synthesizing cDNA, High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, Massachusetts, USA) was used with random primers according to the manufacturer's protocol. The URA3 cDNA was amplified by PCR using semi-nested PCR with three primers (1st, Ura3c_F1 and Ura3c_R1; 2nd, Ura3c_F1 and Ura3c_R2) (Table S1) using Tks Gflex DNA polymerase (Takara Bio, Japan). The product of the second PCR was tailed with adenine using Taq polymerase (Takara Bio, Japan) and cloned into pGEM-T Easy Vector Systems (Promega, Madison, USA). The plasmids were sequenced using ABI 3130 DNA sequencer (Applied Biosystems, California, USA) and using three primers (T7, SP6m and Ura3_R2).

DNA extraction and determining URA3 gene sequence

To determine SSB01 and T01 *URA3* gene sequences, we used the genome sequence of *S. minutum* strain Mf1.05b for reference (Shoguchi *et al.* 2013). As the gene model scaffold311.1|size231230|87274-128405 was predicted to encompass the *URA3* gene, but contain a substantial number of non-sequenced sites, we used another gene model scaffold311.1|size6309|122017-128325, which was predicted mainly by using transcriptome data (Supplementary Fig. S2-3A). We designed specific primers for amplifying the gene by PCR and sequencing (Table S2-1).

We extracted DNA from SSB01 and T01. Cells were collected by centrifugation twice at 5,000 × g for 5 min at 25 °C. Culture medium was discarded, and the cell pellet was re-suspended in 500 µl of MilliQ water (Merck Millipore, Germany). An equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) (Nippon Gene, Japan) and approximately 20 µl of glass beads (Sigma-Aldrich, Merck Millipore, Germany) were added and mixed well by vortexing. After centrifugation, the clear upper aqueous phase containing DNA was carefully transferred to a new microtube and DNA was precipitated by adding an equal volume of isopropanol and 1/10 volume of 3 M sodium acetate. After centrifugation, the DNA pellet was washed with 70% ethanol, dried and dissolved in TE. DNA quality and concentration were checked using a NanoDrop (Thermo Fisher Scientific, Massachusetts, USA). The *URA3* gene fragments were obtained through nested PCR and by using four primers (1st, Ura3g_F1 and Ura3g_R1; 2nd, Ura3g_F2, and Ura3g_R2) using Tks Gflex DNA polymerase (Takara Bio, Japan). The cloned PCR products were sequenced using ABI 3130 DNA sequencer (Applied

Biosystems, California, USA) and using 22 primers (Supplementary Table S2-1, except Ura3g_F1, Ura3g_R1, T7 and SP6m). The nucleotide sequences were deposited at DDBJ/EMBL/GenBank under accession numbers LC363939 and LC363940.

The *URA3* gene sequences were conceptually translated into proteins and used for multiple sequence alignment and phylogenetic analysis as described previously (Nguyen *et al.* 2015). IQ-TREE was used to reconstruct phylogenetic trees using LG + F + G4 model29.

Yeast mutant complementation

Synthetic cDNA sequences encoding *Symbiodinium* URA3 proteins, for which the codon usage was optimized for the budding yeast Saccharomyces cerevisiae, were designed using wild type and T01 mutant strains. The cDNA was digested with restriction enzymes BamHI and PstI (New England Biolabs, Massachusetts, USA) and cloned into the expression vector p414-TEF using T4 ligase (New England Biolabs, Massachusetts, USA) under the control of the TEF1 (Translation Elongation Factor 1-alpha) promoter, which enables strong and constitutive expression. After amplified in E. coli, the vectors were purified and introduced into the yeast strain W303a (MATa ade2-1 his3-11, 15 leu2-3, 112 trp1-1 ura3-1 can1-100) by the lithium acetate method, followed by screening on plates lacking tryptophan. Resulting transformants possessing the empty vector, wild type *URA3* gene and mutant gene were cultured on plates containing or lacking uracil, and examined for complementation of the uracil requiring phenotype.

Symbiosis experiments using the model sea anemone *Exaiptasia pallida*

E. pallida strain H2 was maintained at a density of 3 to 8 animals per plastic case (10 cm diameter), and filled with ASW. Aposymbiotic individuals were kept in dark and the symbiotic ones were maintained under a 12 h light: 12 h dark cycle, with lighting provided at an irradiance of approximately 20 μmol photons m^{-2} s^{-1}, at 25 °C, and fed Clean white shrimp (Kyorin, Japan) 1 to 5 times per week.

Aposymbiotic E. pallida was transferred to a 6 well plate (1 individual/well) and maintained in the psnASWU medium, which was a mixture of psnASW (ASW containing 0.01 mg ml^{-1} of penicillin, 0.01 mg ml^{-1} of streptomycin, 0.02 mg ml^{-1} of neomycin) and 0.2 mg ml^{-1} of uracil, for a day prior to feeding. Aposymbiotic individuals were fed Clean white shrimp mixed with T01 (n = 6), or SSB01 (n = 3) as positive control, or no *Symbiodinium* cells (n = 3) as negative control. After feeding, E. pallida was pre-incubated in psnASWU for 7 days. After pre-incubation, psnAWSU was changed to psnASW. Three T01-fed individuals were kept in psnASWU for comparison. We observed individuals using fluorescence stereo microscopes (Leica, M205FA, Germany) and continued changing psnAWSU or psnAWS daily. Micrographs were processed and analysed using ImageJ software (NIH). For calculating symbiont area, images taken with a Green filter set (to visualize the chlorophyll autofluorescence of any *Symbiodinium* present) were processed by NIH

ImageJ including LPixel ImageJ Plugins (LPX filter2d [filter = bandPassOps__, bpMode = Gaussian, low = 0, hi = 10, postProc = none], freely downloadable from https://lpixel.net/products/lpixel-imagej-plugins/) and threshold set by Otsu algorithm and area measurement was performed by Analyze Particles command (Size = 4-Infinity). Observation was discontinued when the symbiont area was saturated and each symbiont could not be distinguished. The rate of change in symbiont area was calculated as relative values in comparison to the area at the time of medium change from psnAWSU to psnASW. Student's two-tailed t-tests were performed to test whether the area ratios significantly differed between the treatments.

2-3. Results

Screening and phenotyping 5FOA-resistant and nutrient (uracil) -deficient mutants

To select antibiotics and inhibitors suitable for genetic screening experiments, and that could regulate *Symbiodinium* cell growth by their presence or absence, we tested the effects of widely-used antibiotics including kanamycin, neomycin, streptomycin, zeocin, paromomycin and nourseothricin, on cell growth. None of the antibiotics tested had a substantial effect on the algal growth. However, a cell growth inhibitor 5-fluoroorotic acid (5FOA) successfully suppressed the algal growth (Fig. 2-1, WT). 5FOA is a fluorinated derivative of uracil precursor orotic acid and inhibits the growth of cells expressing *URA3* gene, which encodes orotidine-5′-monophosphate (OMP) decarboxylase, through the synthesis of the toxic 5-fluorouracil causing cell death (Supplementary Fig. S2-1).

This inhibitor has also been commonly used in yeast, red alga, and other organisms to isolate 5FOA insensitive, nutrient (uracil)-requiring mutants, i.e. *ura3* mutant (Minoda *et al.* 2004). In addition, only one copy of *URA3* gene in each of the sequenced genomes of *S. minutum* (Shoguchi *et al.* 2013) and *S. microadriaticum* (Aranda *et al.* 2016) was found, suggesting that *URA3* was more promising as a marker gene than multi-gene family members (Supplementary Fig. S2-2A). Thus we used 5FOA for further screening because this inhibitor was useful to establish cell lines whose growth could be regulated by the presence and absence of uracil. For isolating 5FOA-resistant mutants from *Symbiodinium*, we used the strain SSB01 (clade B phylotype), which is an axenic unialgal strain closely related to *S. minutum* Mf1.05b (clade B), the strain whose genome has been sequenced (Xiang *et al.* 2013; Shoguchi *et al.* 2013). We grew SSB01 cells in the nutrient-replete, uracil-containing Marine Broth (MB) medium containing 200 μg/ml 5FOA over eight weeks and successfully isolated four candidate mutant lines. We named these mutant lines T01, T22, T23 and T29, and confirmed that the three of them (T01, T22 and T23) showed stable and reproducible growth suppression in the uracil-limited sea salt medium (IMK) plates and resistance to 5FOA (Fig. 2-1).

Sequence analysis of mutant strains

To identify the mutation sites, we sequenced partial cDNA sequences (from 97th to 532th nucleotide in the 792-bp full length cDNA sequence) of the *URA3* gene of the mutant T01, T22, T23, T29 and wild type SSB01 strains. We found a 9-bp deletion corresponding to three amino acid sites (47-49th) in the T01 cDNA (Fig. 2-2A), while no mutations in the partial cDNA sequences were found in the strains T22, T23 and T29. Homology-based comparison with the crystal structure of the URA3 homolog from other species (Appleby *et al.* 2000) suggested that, although *Symbiodinium* and some other eukaryotes possess homologs only distantly related to well-studied URA3 proteins, this deletion site was proximal to a lysine residue, which is important for substrate recognition, and that the deletion likely affects the enzymatic activity resulting in resistance to 5FOA (Supplementary Fig. S2-2B).

To clarify the cause of the 9-bp deletion in the T01 *URA3* cDNA and search for potential mutation sites in the other mutants, we sequenced a 6-kb genomic DNA (gDNA) fragment containing the *URA3* gene, with reference to a scaffold sequence, scaffold311, and the gene models of the *S. minutum* genome database (Shoguchi *et al.* 2013). We identified a single nucleotide substitution in the intron region of an intron-exon junction in the T01 *URA3* genome (Supplementary Fig. S2-3, Supplementary data), while we found no mutations in the T22, T23 and T29 *URA3* genomic sequences in comparison to the wild type. By sequence comparison using our sequence data from the wild type SSB01 and T01, we found both the canonical intron-exon boundary sequences ('GU-AG') and non-canonical ones within the *URA3* gene, as previously shown in the *S. minutum* genome (Shoguchi *et al.* 2013). The single nucleotide substitution found in this study could alter a non-canonical 'GA-AG(G)' exon-intron junction into 'GA-GG(G)' (Fig. 2-2B), resulting in a splicing variation mutation. To further examine the phenotype-genotype correlations in SSB01 and T01, we cultured them in liquid media in the presence or absence of 5FOA. The growth of T01 was confirmed upon addition of 5FOA (Fig. 2-3), suggesting that the phenotype of T01 was the consequence of mutation(s) affecting uracil synthesis in the pyrimidine biosynthesis pathway (see Discussion and Supplementary Fig. S2-1).

Yeast mutant complementation

Our sequence analyses showed that a single nucleotide substitution was the most likely reason for mis-splicing and the 9-bp deletion in the *URA3* gene at the cDNA level, resulting in 3-amino acid deletion at the protein level (Fig. 2-2). To confirm that the 3-amino acid deletion was responsible for the 5FOA-sensitive and uracil-requiring phenotypes, we conducted yeast complementation assays. For this assay, we cloned the wild type and the mutant (T01) cDNA sequences encoding *Symbiodinium* URA3 proteins with codon usage optimized for the budding yeast Saccharomyces cerevisiae into the vector p414-TEF, transformed a *ura3* mutant yeast cell line with them, and compared the growth on uracil replete and limited media (Fig. 2-4). While the wild type gene rescued the growth of the yeast cells on the uracil-limited medium, the vector control and the T01 gene did not recover the growth of the *ura3* yeast cells.

Symbiosis experiments using the model sea anemone *Exaiptasia pallida*

In order to examine whether the mutant strain T01 maintained the ability to symbiose with host cnidarian animals, we fed the model sea anemone *E. pallida* with T01 and wild type *Symbiodinium* cells. Then we calculated 'area ratio,' which is defined as the ratio of the area having chlorophyll autofluorescence signal of the intracellular algal cells in micrographs taken from the top of the sea anemone at regular time points in comparison to the one at the start of the quantification (day 0), which roughly represented the degree of symbiosis expansion in the host animal body. In the artificial seawater (ASW) medium supplemented with uracil, the area ratio in the *E. pallida* fed with T01 increased over incubation time and was stably sustained, indicating that T01 retained the ability to establish a symbiotic relationship with *E. pallida* in the uracil-replete condition, as the wild type strain did in the uracil-depleted condition (Fig. 2-5). However, when the uracil-containing ASW was replaced with the one lacking uracil, the number of T01 cells living inside the animal host gradually decreased over time, while T01 retained stable symbiosis when the medium was replaced with fresh uracil-containing ASW (Fig. 2-5A, B, Supplementary Fig. S2-4). The area ratio of T01 in the uracil-replete condition increased gradually compared to the wild type in uracil-lacking ASW (Supplementary Fig. S2-4), and the rate of increase in the area ratio seemed to reflect the growth ability of the free-living cells (Figs 2-3 and 2-5C) (see Discussion). These results clearly demonstrate that stable symbiotic relationship between *E. pallida* and T01 was dependent on the availability of uracil in the environment, which is critical for the growth of the algal symbiont (Fig. 2-5C).

2-4. Discussion

Cnidarian-dinoflagellate symbiosis is one of the well-studied and unique model systems that can be used for examining cellular mechanisms of animal-plant symbiosis (Yellowlees *et al.* 2008), but has not been understood fully at the molecular level. The availability of *Symbiodinium* strains possessing conspicuous physiological and/or cellular properties enable easy tracking in symbiosis experiments, and could be an ideal genetic tool. As such no such strain has been available until now. In this study we identified a nutrient-requiring mutant strain of *Symbiodinium* harbouring spontaneous mutations in a gene encoding the uracil synthesis enzyme. We also demonstrated that this mutant can be employed for analysing cellular properties in symbiosis experiments.

It is worth noting that the nutrient-requiring mutants, similar to those screened in this study, serve as potentially useful tools for researchers to develop systems for genetic transformation of *Symbiodinium*. Although gene introduction methods for *Symbiodinium* have been reported(Lohuis and Miller 1998; Ortiz-Matamoros and Villanueva 2015; Ortiz-Matamoros *et al.* 2015), several challenges including low reproducibility, difficulty in isolation, and recovery of actively-growing transformed cells prevent them from being routinely used. The T01 mutant line developed in this study will be a

useful tool to examine whether a gene of interest affects the stability of symbiosis by transformation with a construct containing the target gene and wild type *URA3* either by fusion or as tandemly arranged genes to complement the uracil-requiring phenotype and not by mere transfection of exogenous DNA.

The mutant strains obtained in this study were viable on the medium containing both 5FOA and uracil, which strongly suggests their inability to synthesize uracil due to *URA3* gene mutation and/or suppressed gene expression (Fig. 2-1). Sequencing the cDNA confirmed the 9-bp deletion corresponding to 3-amino acids in the T01 mutant (Fig. 2-2A). By referencing the Bacillus subtilis URA3 protein structure (Appleby *et al.* 2000), the deletion region was predicted to be in the vicinity of a helix containing a lysine residue shown to be important for enzymatic activity (Supplementary Fig. S2-2B). This is consistent with the results of the yeast complementation tests (Fig. 2-4), indicating that the mutant *URA3* gene sufficiently explains the T01 phenotype and that the spontaneous *Symbiodinium URA3* mutant was successfully isolated through the 5-FOA resistance screens. The decreased cell growth rate of T01 compared to wild type strain even in the uracil-replete condition suggested that T01 might have other uncharacterized mutation(s) in the genes associated with growth rate regulation or that a slow growth individual was randomly selected (Fig. 2-3). Once the sexual reproduction cycle of *Symbiodinium* is fully characterized, further sophistication of algal genetic techniques, e.g. mating, backcrossing, as routinely done in the model green alga *Chlamydomonas reinhardtii*, will be useful to segregate mutations associated and not associated with the uracil-requiring phenotype and 'purify' the mutant strain (Rogozin *et al.* 2012).

Sequence comparison of the genomic DNA of the *URA3* gene in T01 and wild type (SSB01) based on the genome database of the reference strain *S. minutum* Mf1.05b (Shoguchi *et al.* 2013) revealed a single nucleotide substitution in an intron (Supplementary Fig. S2-3). A previous genome study (Shoguchi *et al.* 2013) suggested that many *S. minutum* genes, including *URA3*, possess divergent atypical exon-intron boundary structures, and that the canonical splicing donor site (GU) and acceptor site (AG) sequences were not necessarily conserved. The intron junction sequence where we identified the nucleotide substitution did not follow the 'GU-AG' rule, but was found to be 'GA-AG' in the wild type strain. Interestingly, the mutation in T01 had substituted the acceptor 'AG' with 'GG,' resulting in defective mRNA splicing. The new acceptor was not the first 'AG' that was positioned downstream to the original acceptor but the second 'AG' (Fig. 2-2B). The original and second downstream 'AG' were followed by a 'G,' in contrast to the first downstream 'AG' followed by an 'A,' indicating that the splicing junction was probably recognized as 'GA-AG-G' including the first nucleotide of the downstream exon. To our knowledge, this is the first study using a splicing variant mutant of *Symbiodinium*, illustrating the unusual exon-intron boundary recognition functioning in vivo, as had been predicted in the genome analysis (Shoguchi *et al.* 2013).

Our results invoke further questions on the evolution and regulation of such uncanonical splicing mechanisms.

First, the mechanisms of junction site recognition and determination remain unknown; e.g. the potentially recognizable accepter 'AGG' sequence was also located upstream to the original site but was not recognized for splicing (Shoguchi *et al.* 2013). Second, it is still unclear whether the nucleotide sequences are sufficient for determining the junctions or if other factors such as spliceosomal RNA and proteins are involved (Parkinson *et al.* 2016). Recent advances in high throughput sequencing technology may be helpful in tackling these issues. Accumulating large amounts of transcriptomic and proteomic data from *Symbiodinium* culture strains and environmental samples can be useful in identifying natural variations in splicing junctions in conserved proteins. This will enable us to estimate how frequently acceptable protein sequence alterations resulting from splicing variations occur (Bieri *et al.* 2016; Weis *et al.* 2018). Biochemical analysis of dinoflagellate spliceosomes as well as genetic transformation system using the *Symbiodinium* mutant strains developed in this study will also be of great assistance in understanding the evolution of such complex and unusual splicing mechanisms in dinoflagellates.

Our co-culture experiments showed that T01 was able to maintain a stable symbiotic relationship with the model sea anemone E. pallida in uracil-containing ASW, as well as the wild type strain (Fig. 2-5), indicating that in T01 the cellular machinery involved in the symbiosis was not impaired. However, the symbiotic status between *E. pallida* and T01 became unstable when it was cultured in uracil-free medium (Fig. 2-5A and B), suggesting that the availability of uracil in the medium was a requisite for sustaining the stable symbiosis. Ten days after depletion of uracil, the symbiosed T01 cells were only sparsely distributed in the anemone body (Fig. 2-5A), which appeared to mimic the 'bleaching' status (Baghdasarian and Muscatine 2000; Bieri *et al.* 2016). The unsuccessful symbiosis in the uracil-depleted condition suggested that the supply of uracil from the host to the symbiont was not enough, if any, to sustain the proliferation of the symbiont. Further, with the use of the mutant T01, it is now possible to experimentally 'switch on and off' the sea anemone-algal symbiosis by using media containing or lacking uracil, respectively. This has important implications on the relationship between symbiosis stability and the growth ability of the symbiont cell. In the free-living condition, T01 cell cultures showed the significantly increased cell growth depending on the availability of uracil 28 days after the onset of the medium change (Fig. 2-5C). Although it is difficult to directly compare the symbiotic and free-living conditions, the uracil-dependent cell proliferation in the free-living condition can explain to some extent how the symbiosis was established in the uracil-dependent manner. A plausible interpretation is that a certain level of cell proliferation is necessary for sustaining symbiosis. Considering that dividing *Symbiodinium* cells were preferentially expelled from cnidarian hosts (Hoegh-Guldberg 1999), even though a certain number of cells are expelled, their daughter cells can be re-symbiosed with the host after cell division resulting in the expansion of symbiosis if they outnumber the originally expelled cells. In the case of uracil-depleted T01 unable to proliferate without uracil, the number of cells re-entering the host endodermal cells

decrease, thereby, leading to loss of algae inside the animal host (Fig. 2-5).

It should be noted that, although uracil was not supplied, or if any very limited, from the host to the symbiont in the model sea anemone E. pallida, other cnidarian hosts including corals may have different metabolic properties and are to be examined in future studies. This suggests that, although cautions are needed in interpreting results under less controlled experimental conditions, e.g. in the field or aquarium tank, the mutant strains developed in this study would be useful for studying metabolic interactions between hosts and symbionts, and also for screening host cnidarian species which supply uracil and possibly other basic metabolites to symbionts.

Cnidarian-algal endosymbiosis has been an important study model in ecology, genomics and cell biology due to its huge impact on marine ecosystems, especially in tropical and subtropical areas (Davy *et al.* 2012). Previous studies have shown that elevated sea water temperature could lead to the collapse of symbiosis and coral 'bleaching' (Hoegh-Guldberg *et al.* 2007; Maruyama *et al.* 2011; Hughes *et al.* 2017). Thus, understanding the mechanisms of maintaining stability of symbiosis is key to predicting possible effects of environmental changes on marine ecosystems. To our knowledge, the first mutant strain of *Symbiodinium* established in this study emphasizes the importance of cell proliferation in sustaining the symbiosis in vivo, and can be used to investigate the molecular mechanisms of the symbiosis in future studies. This will be a powerful tool for *Symbiodinium* genetics research, and for advancing 'symbiotic genetics', through which it is possible to examine what kinds of genes are relevant for establishing stable symbiotic relationships with cnidarian hosts and for using genetically engineered symbiotic algae.

Figures 2

Figure 2-1. Phenotypes of *Symbiodinium* wild type and mutant cells. Wild type (WT) and mutant (T01, T22 and T23) cells were spotted and grown on agar plates with complete medium (MB), which originally contains uracil, with or without 5FOA, and the minimal medium (IMK) with or without uracil.

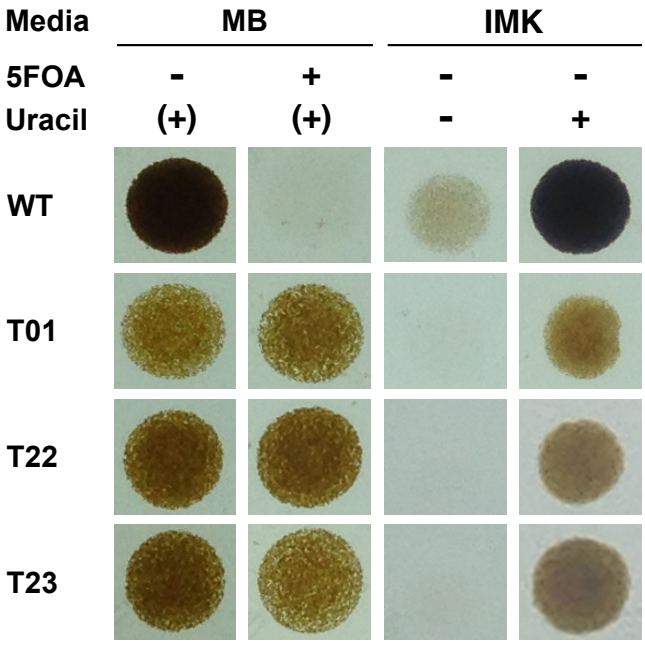


Figure 2-2. Sequences of *URA3* cDNA and splicing variation on mRNA between wild type and T01. (**A**) A deletion of 9 bp, corresponding to 3 amino acids (47-49th), found in the T01 *URA3* cDNA is indicated in a black closed box. (**B**) Grey boxes and bold letters represent introns and nucleotides relevant to intron-exon junction recognition, respectively. Letters with bars represent the nucleotide site where the mutation is mapped in the T01 genome.

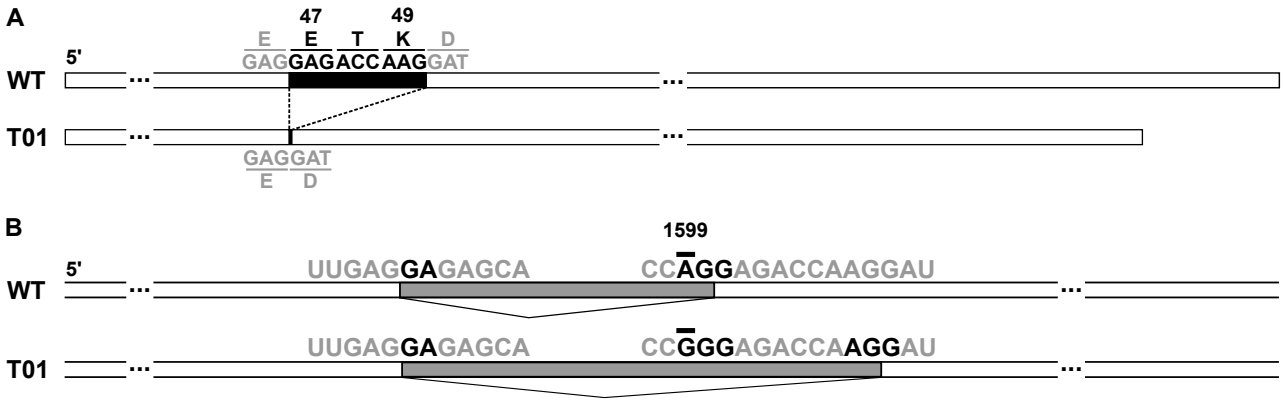


Figure 2-3. Effects of 5FOA on the cell growth rates of wild type and T01 in MB liquid cultures. WT (square) and T01 (circle) cells were grown in the absence (closed symbol) or presence (open symbol) of 5FOA (*adjusted p for interaction <0.05, n = 3, *t*-test).

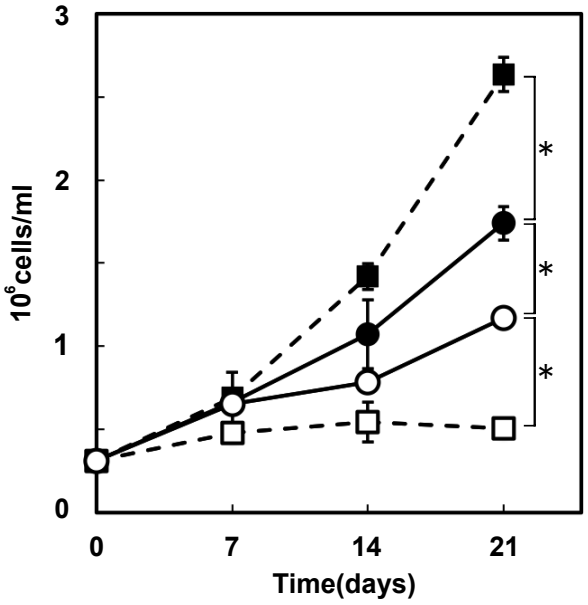


Figure 2-4. Functional complementation of *Symbiodinium URA3* genes in a yeast *ura3* mutant. The growth of the yeast *ura3* mutant cells transformed with an empty vector (vector), the wild type *URA3* cDNA (WT) and the mutant cDNA (T01) on plate containing (+Uracil) or lacking uracil (−Uracil).

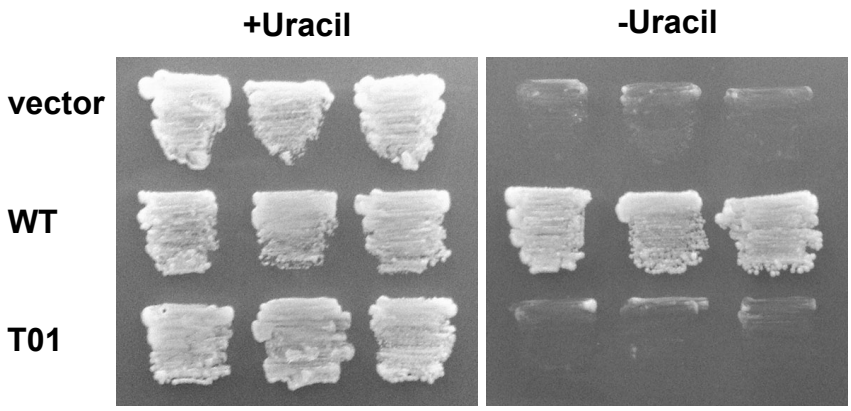
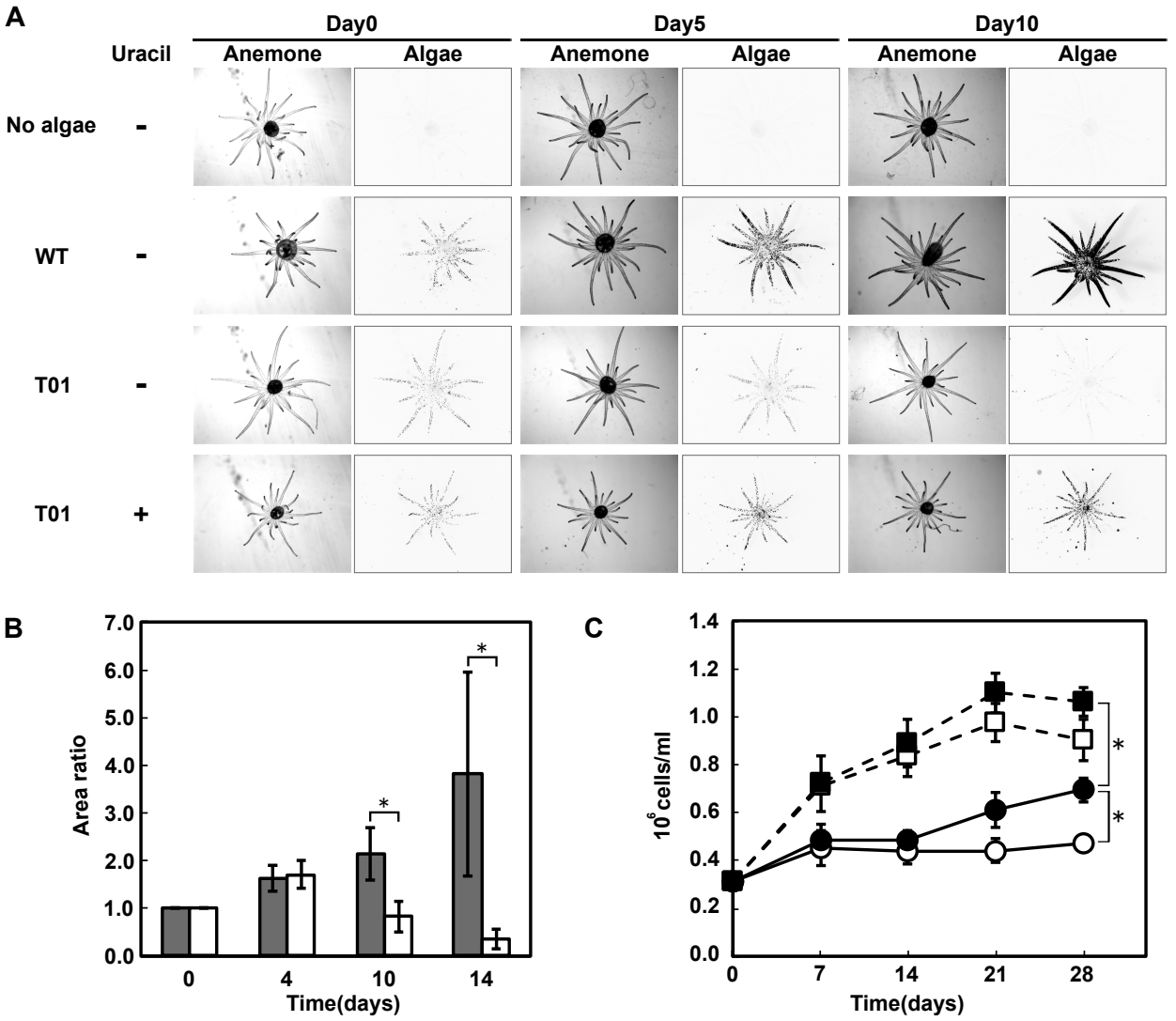


Figure 2-5. Symbiosis experiments using wild type and T01 cells. (**A**) Stereomicrographs using bright field (Anemone) or a green excitation filter set (Algae: chlorophyll autofluorescence of *Symbiodinium*) photographed with time intervals shown above. Representative images of individuals are shown for each treatment. (**B**) Effect of uracil on the symbiotic state of T01. *Symbiodinium* area ratios were quantified by comparing signal regions of chlorophyll and normalized to the value of day 0 in the presence (closed bar) or absence (open bar) of uracil (*$p < 0.05$, $n = 3$, *t*-test). (**C**) Effects of uracil on the cell growth rate of WT and T01 in ASW liquid culture. WT (square) and T01 (circle) cells were grown in the presence (closed symbol) or absence (open symbol) of uracil (*adjusted p for interaction <0.05, $n = 3$, *t*-test).



General discussion

Symbiosis is a two-edged blade: it could produce great benefits in a certain condition, especially in competitive or low nutrient environments, while its dependency to the partner species also could be risks and constraints for survival in other situations. Cnidarian animals such as corals and sea anemones have highly established symbiotic relationship with Symbiodiniaceae while they often 'break up' under stressful environment changes. One hypothesis is that bleaching is an emergency workaround to avoid fetal damage to the host since symbiotic state could be toxic under stress (Weis 2008). Considering the results in Chapter 1 showing that uracil free treatment collapsed symbiosis between *E. diaphana* and T01, non-proliferating symbionts may also be involved in the early process of bleaching. Rather, I hypothesize that the decreased proliferation of symbionts can be a triggering factor of bleaching under environment stress. Many cultured Symbiodiniaceae strains decreased photosynthetic activity and proliferation under 33℃ heat stress, where *E. diaphana* and many symbiotic cnidarians normally cause bleaching (Karim *et al.* 2015). Interestingly, the decrease in photosynthetic activity was always associated with the decrease in proliferation, but not the other way around. This result suggests that symbionts' proliferation might be more susceptible to heat stress than photosynthetic activity, and more substantial influence on bleaching. In Chapter 2, T01 provided important insights: Symbiosis between *E. diaphana* and T01 was established in uracil containing normal seawater whereas it was collapsed when the seawater was substituted to the uracil-free one. Importantly, T01 was just unable to proliferate but not dead under uracil-free condition. These results suggest that a metabolic change in symbiont could lead collapse of symbiosis. A future study to test the causalities of proliferation on breaching under other types of environment changes should be conducted.

A remaining question is in what way the host *E. diaphana* can recognize how Symbiodiniaceae symbionts are doing within their cells. In Chapter 1, symbiotic state specific-heat stress induced genes were screened in *E. diaphana*, and genes involved in carbohydrate and protein metabolisms in lysosomes were detected. In Symbiodiniaceae, genes involved in heat shock response, calcium signaling, organellar protein transport, and sugar metabolism were detected. One plausible hypothesis is that the HIBA genes play key roles in lysosomal (or symbiosomal) degradation and modification of glycoproteins at the symbiont cell surface (Winchester 2005) and thereby affecting the symbiosis stability under heat stress. Glycoproteins on the cell surface of symbiont are considered to play important roles in the recognition of symbionts by the host coral (Takeuchi *et al.* 2017; Huang *et al.* 2017).

In Symbiodiniaceae, a nucleotide-sugar transporter *GONST3* gene was screened. *GONST3* is thought to function in the import of nucleotide-sugar from cytosol to the Golgi apparatus for downstream glycosylation reactions. As mentioned earlier, it is suggested that sugar, more specifically glycoproteins, is an important for the recognition of symbionts by the host (Takeuchi *et al.* 2017; Huang *et al.* 2017). Given these findings, I hypothesize that cytosolic sugar

metabolism and Golgi apparatus-mediated glycosylation of proteins and/or cell wall components can be susceptible to stress and damage when symbionts are exposed to heat in *hospite*. The host may recognize 'trouble' or dysfunction in symbiont under heat stress by detecting damaged glycoproteins on the surface of symbiont with uncharacterized mechanisms. In other words, symbiont may present own normality or productivity with the surface glycoproteins while the host may monitor it by metabolizing the glycoproteins. This hypothesis will be able to be tested by using presently developed Symbiodiniaceae mutants (Chapter 2) and future gene induction methods.

The present study succeeded to establish Symbiodiniaceae uracil auxotroph mutant strain 'T01'. This is a breakthrough to open a way for the development of gene induction methods in Symbiodiniaceae. Although various gene induction methods have been developed in other species, only quite a few studies have been reported in Symbiodiniaceae. ten Lohuis and Miller proposed a mixing of silicon carbide whisker and induced genes in 1998 (Lohuis and Miller 1998). Unfortunately, no report reproducing it has been emerged since then. Recently, Ortis-Matamoros's group proposed two new methods, glass beads/PEG method (Ortiz-Matamoros and Villanueva 2015), and Agrobacterium method (Ortiz-Matamoros *et al.* 2015). However, these methods have not been reproduced by other groups so far and are transient gene induction so that it cannot use for establishing transformant strains that possesses stable expression of introduced gene. In general, difficulty in developing gene introduction methods in a non-model organism relies on how efficient available methods to screen transfected cells are. To develop a gene introduction method from the beginning, a lot of parameters and conditions need to be tested. The present uracil auxotroph strain T01 has a point mutation in *URA3* which is essential for uracil synthesis so that it is unable to proliferate in the absence of uracil. Introducing a *URA3* gene and successfully transforming cells is only way for the algae to grow in the absence of uracil. This simple screening system using T01 can be a strong tool to streamline such enormous trials for setting up novel transformation methods.

References

Anders, S., P. T. Pyl, and W. Huber, 2015 HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169.

Appleby, T. C., C. Kinsland, T. P. Begley, and S. E. Ealick, 2000 The crystal structure and mechanism of orotidine 5'-monophosphate decarboxylase. Proc. Natl. Acad. Sci. U. S. A. 97: 2005–2010.

Aranda, M., Y. Li, Y. J. Liew, S. Baumgarten, O. Simakov *et al.*, 2016 Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. Sci. Rep. 6: 39734.

Baghdasarian, G., and L. Muscatine, 2000 Preferential expulsion of dividing algal cells as a mechanism for regulating algal-cnidarian symbiosis. Biol. Bull. 199: 278–286.

Bannai, H., Y. Tamada, O. Maruyama, K. Nakai, and S. Miyano, 2002 Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18: 298–305.

Barott, K. L., A. A. Venn, S. O. Perez, S. Tambutté, and M. Tresguerres, 2015 Coral host cells acidify symbiotic algal microenvironment to promote photosynthesis. Proc. Natl. Acad. Sci. U.S.A. 112: 607–612.

Barshis, D. J., J. T. Ladner, T. A. Oliver, and S. R. Palumbi, 2014 Lineage-specific transcriptional profiles of *Symbiodinium* spp. unaltered by heat stress in a coral host. Mol. Biol. Evol. 31: 1343–1352.

Baumgarten, S., O. Simakov, L. Y. Esherick, Y. J. Liew, E. M. Lehnert *et al.*, 2015 The genome of *Aiptasia*, a sea anemone model for coral symbiosis. Proc. Natl. Acad. Sci. U.S.A. 112: 11893–11898.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: A Practical and powerful approach to multiple testing. J. Roy. Statist. Soc. 57: 289–300.

Bieri, T., M. Onishi, T. Xiang, A. R. Grossman, and J. R. Pringle, 2016 Relative contributions of various cellular mechanisms to loss of algae during cnidarian bleaching. PLoS ONE 11: e0152693.

Biquand, E., N. Okubo, Y. Aihara, V. Rolland, D. C. Hayward *et al.*, 2017 Acceptable symbiont cell size differs among cnidarian species and may limit symbiont diversity. The ISME Journal 11: 1702–1712.

Brown, B. E., 1997 Coral bleaching: causes and consequences. Coral Reefs 16: S129–S138.

Burriesci, M. S., T. K. Raab, and J. R. Pringle, 2012 Evidence that glucose is the major transferred metabolite in dinoflagellate-cnidarian symbiosis. J. Exp. Biol. 215: 3467–3477.

Chen, M.-C., M.-C. Hong, Y.-S. Huang, M.-C. Liu, Y.-M. Cheng *et al.*, 2005 ApRab11, a cnidarian homologue of the recycling regulatory protein Rab11, is involved in the establishment and maintenance of the *Aiptasia-Symbiodinium* endosymbiosis. Biochem. Biophys. Res. Commun. 338: 1607–1616.

Cleves, P. A., M. E. Strader, L. K. Bay, J. R. Pringle, and M. V. Matz, 2018 CRISPR/Cas9-mediated genome editing in a reef-building coral. Proc. Natl. Acad. Sci. U. S. A. 115: 201722151–5240.

Cumbo, V. R., A. H. Baird, and M. J. H. Van OPPEN, 2013 The promiscuous larvae: flexibility in the establishment of symbiosis in corals. Coral Reefs 32: 111–120.

Daly, M., and D. Fautin, 2018 World List of Actiniaria. Exaiptasia diaphana (Rapp, 1829). Accessed through: World Register of Marine Species at: http://www.marinespecies.org/aphia.php?p=taxdetails&id=1264073 on 2018-12-15

Dani, V., P. Ganot, F. Priouzeau, P. Furla, and C. Sabourault, 2014 Are Niemann-Pick type C proteins key players in cnidarian-dinoflagellate endosymbioses? Mol. Ecol. 23: 4527–4540.

Davy, S. K., D. Allemand, and V. M. Weis, 2012 Cell biology of cnidarian-dinoflagellate symbiosis. Microbiol. Mol. Biol. Rev. 76: 229–261.

Desalvo, M. K., S. Sunagawa, C. R. Voolstra, and M. Medina, 2010 Transcriptomic responses to heat stress and bleaching in the elkhorn coral *Acropora palmata*. Mar. Ecol. Prog. Ser. 402: 97–113.

Desalvo, M. K., C. R. Voolstra, S. Sunagawa, J. A. Schwarz, J. H. Stillman *et al.*, 2008 Differential gene expression during thermal stress and bleaching in the Caribbean coral Montastraea faveolata. Mol. Ecol. 17: 3952–3971.

Dunn, S. R., J. C. Thomason, M. D. A. Le Tissier, and J. C. Bythell, 2004 Heat stress induces different forms of cell death in sea anemones and their endosymbiotic algae depending on temperature and duration. Cell Death Differ. 11: 1213–1222.

Ganot, P., A. Moya, V. Magnone, D. Allemand, P. Furla *et al.*, 2011 Adaptations to endosymbiosis in a cnidarian-dinoflagellate association: differential gene expression and specific gene duplications (E. Rulifson, Ed.). PLoS

Genet. 7: e1002187.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29: 644–652.

Grajales, A., and E. Rodríguez, 2014 Morphological revision of the genus *Aiptasia* and the family Aiptasiidae (Cnidaria, Actiniaria, Metridioidea). Zootaxa 3826: 55–100.

Hambleton, E. A., A. Guse, and J. R. Pringle, 2014 Similar specificities of symbiont uptake by adults and larvae in an anemone model system for coral biology. J. Exp. Biol. 217: 1613–1619.

Hawkins, T. D., B. J. Bradley, and S. K. Davy, 2013 Nitric oxide mediates coral bleaching through an apoptotic-like cell death pathway: evidence from a model sea anemone-dinoflagellate symbiosis. FASEB J. 27: 4790–4798.

Hillyer, K. E., S. Tumanov, S. Villas-Bôas, and S. K. Davy, 2016 Metabolite profiling of symbiont and host during thermal stress and bleaching in a model cnidarian-dinoflagellate symbiosis. J. Exp. Biol. 219: 516–527.

Hoegh-Guldberg, O., 1999 Climate change, coral bleaching and the future of the world's coral reefs. Mar. Freshwater Res. 50: 839–866.

Hoegh-Guldberg, O., P. J. Mumby, A. J. Hooten, R. S. Steneck, P. Greenfield *et al.*, 2007 Coral reefs under rapid climate change and ocean acidification. Science 318: 1737–1742.

Hohman, T. C., P. L. McNEIL, and L. Muscatine, 1982 Phagosome-lysosome fusion inhibited by algal symbionts of *Hydra viridis*. J. Cell Biol. 94: 56–63.

Hong, M.-C., Y.-S. Huang, P.-C. Song, W.-W. Lin, L.-S. Fang *et al.*, 2009 Cloning and characterization of ApRab4, a recycling Rab protein of *Aiptasia pulchella*, and its implication in the symbiosome biogenesis. Mar. Biotechnol. 11: 771–785.

Houlbrèque, F., and C. F. Pagès, 2009 Heterotrophy in Tropical Scleractinian Corals. Biol. Rev. Camb. Philos. Soc. 84: 1–17.

Huang, K.-J., Z.-Y. Huang, C.-Y. Lin, L.-H. Wang, P.-H. Chou *et al.*, 2017 Generation of clade- and symbiont-specific antibodies to characterize marker molecules during Cnidaria-*Symbiodinium* endosymbiosis. Sci. Rep. 7: 5488.

Hughes, T. P., J. T. Kerry, M. Álvarez-Noriega, J. G. Álvarez-Romero, K. D. Anderson *et al.*, 2017 Global warming and recurrent mass bleaching of corals. Nature 543: 373–377.

Jones, V. A. S., M. Bucher, E. A. Hambleton, and A. Guse, 2018 Microinjection to deliver protein, mRNA, and DNA into zygotes of the cnidarian endosymbiosis model Aiptasia sp. Sci. Rep. 8: 16437.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin, 2017 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14: 587–589.

Karim, W., S. Nakaema, and M. Hidaka, 2015 Temperature Effects on the Growth Rates and Photosynthetic Activities of Symbiodinium Cells. Journal of Marine Science and Engineering 2018, Vol. 6, Page 13 3: 368–381.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley *et al.*, 2013 TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14: R36.

Kuo, J., M.-C. Chen, C.-H. Lin, and L.-S. Fang, 2004 Comparative gene expression in the symbiotic and aposymbiotic *Aiptasia pulchella* by expressed sequence tag analysis. Biochem. Biophys. Res. Commun. 318: 176–186.

LaJeunesse, T. C., J. E. Parkinson, P. W. Gabrielson, H. J. Jeong, J. D. Reimer *et al.*, 2018 Systematic Revision of Symbiodiniaceae Highlights the Antiquity and Diversity of Coral Endosymbionts. Current Biology 28: 2570–2580.e6.

Lehnert, E. M., M. S. Burriesci, and J. R. Pringle, 2012 Developing the anemone Aiptasia as a tractable model for cnidarian-dinoflagellate symbiosis: the transcriptome of aposymbiotic A. pallida. BMC Genomics 13: 271.

Lehnert, E. M., M. E. Mouchka, M. S. Burriesci, N. D. Gallo, J. A. Schwarz *et al.*, 2014 Extensive Differences in Gene Expression Between Symbiotic and Aposymbiotic Cnidarians. G3: Genes|Genomes|Genetics 4: 277–295.

Lin, S., S. Cheng, B. Song, X. Zhong, X. Lin *et al.*, 2015 The Symbiodinium kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis. Science 350: 691–694.

Liu, H., T. G. Stephens, R. A. González-Pech, V. H. Beltran, B. Lapeyre *et al.*, 2018 Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. Commun. Biol. 2018 1:1 1: 95.

Lohuis, ten, M. R., and D. J. Miller, 1998 Genetic transformation of dinoflagellates (*Amphidinium* and *Symbiodinium*):

expression of GUS in microalgae using heterologous promoter constructs. Plant J. 13: 427–435.

Maruyama, S., T. Suzaki, A. P. M. Weber, J. M. Archibald, and H. Nozaki, 2011 Eukaryote-to-eukaryote gene transfer gives rise to genome mosaicism in euglenids. BMC Evol. Biol. 11: 105.

Mastorodemos, V. M., D. K. Kotzamani, I. Z. Zaganas, G. A. Arianoglou, H. L. Latsoudis *et al.*, 2009 Human GLUD1 and GLUD2 glutamate dehydrogenase localize to mitochondria and endoplasmic reticulum. Biochemistry and Cell Biol. 87: 505–516.

Matthews, J. L., C. M. Crowder, C. A. Oakley, A. Lutz, U. Roessner *et al.*, 2017 Optimal nutrient exchange and immune responses operate in partner specificity in the cnidarian-dinoflagellate symbiosis. Proc. Natl. Acad. Sci. U.S.A. 114: 201710733–13199.

Meiri, D., and A. Breiman, 2009 Arabidopsis ROF1 (FKBP62) modulates thermotolerance by interacting with HSP90.1 and affecting the accumulation of HsfA2-regulated sHSPs. Plant J. 59: 387–399.

Michaeli, S., and H. Fromm, 2015 Closing the loop on the GABA shunt in plants: are GABA metabolism and signaling entwined? Front. Plant. Sci. 6: 419.

Mies, M., C. R. Voolstra, C. B. Castro, D. O. Pires, E. N. Calderon *et al.*, 2017 Expression of a symbiosis-specific gene in Symbiodinium type A1 associated with coral, nudibranch and giant clam larvae. R Soc Open Sci 4: 170253.

Minoda, A., R. Sakagami, F. Yagisawa, T. Kuroiwa, and K. Tanaka, 2004 Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. Plant Cell Physiol. 45: 667–671.

Mitchelmore, C. L., E. A. Verde, A. H. Ringwood, and V. M. Weis, 2003 Differential accumulation of heavy metals in the sea anemone Anthopleura elegantissima as a function of symbiotic state. Aquatic Toxicol. 64: 317–329.

Mohamed, A. R., V. Cumbo, S. Harii, C. Shinzato, C. X. Chan *et al.*, 2016 The transcriptomic response of the coral Acropora digitifera to a competent Symbiodinium strain: the symbiosome as an arrested early phagosome. Mol. Ecol. 25: 3127–3141.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold, 2008 Mapping and quantifying mammalian

transcriptomes by RNA-Seq. Nat. Methods 5: 621–628.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh, 2015 IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32: 268–274.

Oakley, C. A., M. F. Ameismeier, L. Peng, V. M. Weis, A. R. Grossman *et al.*, 2016 Symbiosis induces widespread changes in the proteome of the model cnidarian *Aiptasia*. Cell Microbiol. 18: 1009–1023.

Oakley, C. A., E. Durand, S. P. Wilkinson, L. Peng, V. M. Weis *et al.*, 2017 Thermal shock induces host proteostasis disruption and endoplasmic reticulum stress in the model symbiotic cnidarian *Aiptasia*. J. Proteome Res. 16: 2121–2134.

Ortiz-Matamoros, M. F., and M. A. Villanueva, 2015 Transient transformation of cultured photosynthetic dinoflagellates (Symbiodinium spp.) with plant-targeted vectors. Cien. Mar. 41(1): 21-32

Ortiz-Matamoros, M. F., T. Islas-Flores, B. Voigt, D. Menzel, F. Baluška *et al.*, 2015 Heterologous DNA Uptake in Cultured Symbiodinium spp. Aided by Agrobacterium tumefaciens. PLoS ONE 10: e0132693–16.

Parkinson, J. E., S. Baumgarten, C. T. Michell, I. B. Baums, T. C. LaJeunesse *et al.*, 2016 Gene Expression Variation Resolves Species and Individual Strains among Coral-Associated Dinoflagellates within the Genus Symbiodinium. Genome Biol. Evol. 8: 665–680.

Pernice, M., A. Meibom, A. Van Den Heuvel, C. Kopp, I. Domart-Coulon *et al.*, 2012 A single-cell view of ammonium assimilation in coral–dinoflagellate symbiosis. The ISME Journal 6: 1314–1324.

Picard, D., 2002 Heat-shock protein 90, a chaperone for folding and regulation. CMLS, Cell. Mol. Life Sci. 59: 1640–1648.

Pierleoni, A., V. Indio, C. Savojardo, P. Fariselli, P. L. Martelli *et al.*, 2011 MemPype: a pipeline for the annotation of eukaryotic membrane proteins. Nucleic Acids Res. 39: W375–W380.

Pochon, X., and R. D. Gates, 2010 A new Symbiodinium clade (Dinophyceae) from soritid foraminifera in Hawai'i. Mol. Phylogenet. Evol. 56: 492–497.

Richier, S., M. Rodriguez-Lanetty, C. E. Schnitzler, and V. M. Weis, 2008 Response of the symbiotic cnidarian

Anthopleura elegantissima transcriptome to temperature and UV increase. Comparative Biochemistry and Physiology - Part D 3: 283–289.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth, 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.

Rodriguez-Lanetty, M., W. S. Phillips, and V. M. Weis, 2006 Transcriptome analysis of a cnidarian – dinoflagellate mutualism reveals complex modulation of host gene expression. BMC Genomics 7: 23.

Rogozin, I. B., L. Carmel, M. Csuros, and E. V. Koonin, 2012 Origin and evolution of spliceosomal introns. Biol. Direct 7: 11.

Rosic, N. N., M. Pernice, S. Dove, S. R. Dunn, and O. Hoegh-Guldberg, 2011 Gene expression profiles of cytosolic heat shock proteins Hsp70 and Hsp90 from symbiotic dinoflagellates in response to thermal stress: possible implications for coral bleaching. Cell Stress Chaperones 16: 69–80.

Schwartz, S. L., C. Cao, O. Pylypenko, A. Rak, and A. Wandinger-Ness, 2007 Rab GTPases at a glance. J. Cell. Sci. 120: 3905–3910.

Seto, S., K. Tsujimura, and Y. Koide, 2011 Rab GTPases Regulating Phagosome Maturation Are Differentially Recruited to Mycobacterial Phagosomes. Traffic 12: 407–420.

Shoguchi, E., G. Beedessee, I. Tada, K. Hisata, T. Kawashima *et al.*, 2018 Two divergent Symbiodinium genomes reveal conservation of a gene cluster for sunscreen biosynthesis and recently lost genes. BMC Genomics 19: 458.

Shoguchi, E., C. Shinzato, T. Kawashima, F. Gyoja, S. Mungpakdee *et al.*, 2013 Draft assembly of the Symbiodinium minutum nuclear genome reveals dinoflagellate gene structure. Curr. Biol. 23: 1399–1408.

Sproles, A. E., N. L. Kirk, S. A. Kitchen, C. A. Oakley, A. R. Grossman *et al.*, 2018 Phylogenetic characterization of transporter proteins in the cnidarian-dinoflagellate symbiosis. Molecular Phylogenet. Evol. 120: 307–320.

Stat, M., D. Carter, and O. Hoegh-Guldberg, 2006 The evolutionary history of *Symbiodinium* and scleractinian hosts-symbiosis, diversity, and the effect of climate change. Perspectives in Plant Ecol.

Steele, R. D., 1977 The Significance of Zooxanthella-Containing Pellets Extruded by Sea Anemones. Bulletin of Marine

Science 27: 591–594.

Sun, J., T. Nishiyama, K. Shimizu, and K. Kadota, 2013 TCC: an R package for comparing tag count data with robust normalization strategies. BMC Bioinformatics 14: 219.

Sunagawa, S., E. C. Wilson, M. Thaler, M. L. Smith, C. Caruso *et al.*, 2009 Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. BMC Genomics 10: 258.

Takahashi, S., S. Whitney, S. Itoh, T. Maruyama, and M. Badger, 2008 Heat stress causes inhibition of the de novo synthesis of antenna proteins and photobleaching in cultured Symbiodinium. Proc. Natl. Acad. Sci. U. S. A. 105: 4203–4208.

Takeuchi, R., M. Jimbo, F. Tanimoto, C. Tanaka, S. Harii *et al.*, 2017 Establishment of a model for chemoattraction of *Symbiodinium* and characterization of chemotactic compounds in *Acropora tenuis*. Fisheries Science 83: 479–487.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan *et al.*, 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28: 511–515.

Wakefiel, T. S., and S. C. Kempf, 2001 Development of host- and symbiont-specific monoclonal antibodies and confirmation of the origin of the symbiosome membrane in a cnidarian-dinoflagellate symbiosis. Biol. Bull. 200: 127–143.

Weis, V. M., 2008 Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. J Exp Biol 211: 3059–3066.

Weis, V. M., S. K. Davy, O. Hoegh-Guldberg, M. Rodriguez-Lanetty, and J. R. Pringle, 2018 Cell biology in model systems as the key to understanding corals. Trends Ecol. Evol. (Amst.) 23: 369–376.

Winchester, B., 2005 Lysosomal metabolism of glycoproteins. Glycobiology 15: 1R–15R.

Xiang, T., E. A. Hambleton, J. C. DeNofrio, J. R. Pringle, and A. R. Grossman, 2013 Isolation of clonal axenic strains of the symbiotic dinoflagellate Symbiodinium and their growth and host specificity. J. Phycol. 49: 447–458.

Yamashita, H., and K. Koike, 2011 Search for clues to clarify the mechanism of coral-zooxanthella symbiosis, from the studies inside and outside corals. Midoriishi: report of Akajima Marine Sci. Lab. (22):14-20. (in Japanese)

Yellowlees, D., T. A. V. Rees, and W. Leggat, 2008 Metabolic interactions between algal symbionts and invertebrate hosts. Plant Cell Environ. 31: 679–694.

Young, M. D., M. J. Wakefield, G. K. Smyth, and A. Oshlack, 2010 Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11: R14.

Supplementary tables and figures

Supplementary tables and figures 1

Figure S1-1. Phylogenetic tree of 28S LSU rDNAs

A maximum likelihood tree of 28S LSU rDNAs from the family Symbiodiniaceae and the related dinoflagellates is shown. Support values using an SH-like approximate likelihood ratio test (left) and standard bootstrap support (right) are shown on each branch.

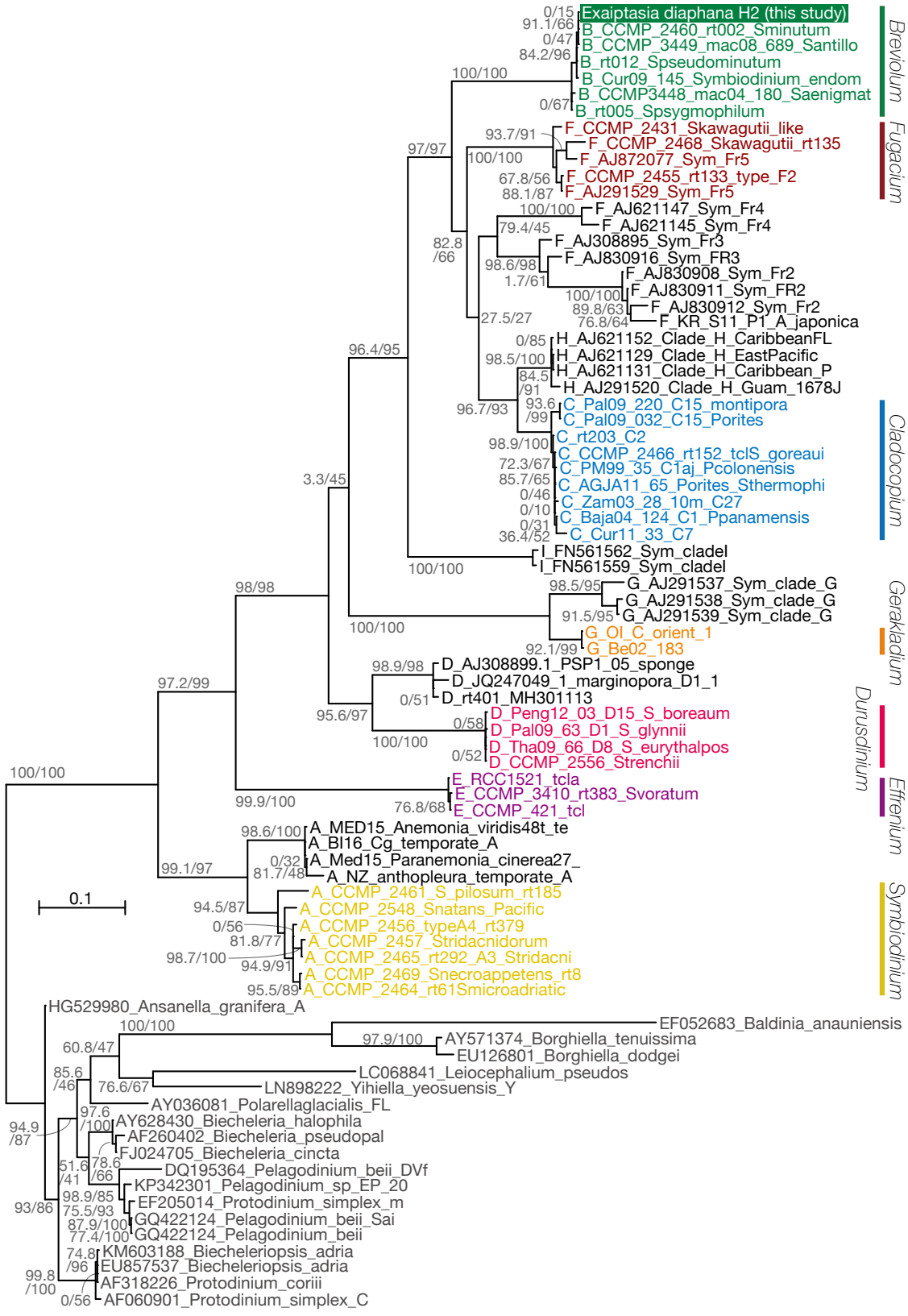

Exaiptasia diaphana H2 (this study)
B_CCMP_2460_rt002_Sminutum
B_CCMP_3449_mac08_689_Santillo
B_rt012_Spseudominutum
B_Cur09_145_Symbiodinium_endom
B_CCMP3448_mac04_180_Saenigmat
B_rt005_Spsygmophilum
F_CCMP_2431_Skawagutii_like
F_CCMP_2468_Skawagutii_rt135
F_AJ872077_Sym_Fr5
F_CCMP_2455_rt133_type_F2
F_AJ291529_Sym_Fr5
F_AJ621147_Sym_Fr4
F_AJ621145_Sym_Fr4
F_AJ308895_Sym_Fr3
F_AJ830916_Sym_FR3
F_AJ830908_Sym_Fr2
F_AJ830911_Sym_FR2
F_AJ830912_Sym_Fr2
F_KR_S11_P1_A_japonica
H_AJ621152_Clade_H_CaribbeanFL
H_AJ621129_Clade_H_EastPacific
H_AJ621131_Clade_H_Caribbean_P
H_AJ291520_Clade_H_Guam_1678J
C_Pal09_220_C15_montipora
C_Pal09_032_C15_Porites
C_rt203_C2
C_CCMP_2466_rt152_tclS_goreaui
C_PM99_35_C1aj_Pcolonensis
C_AGJA11_65_Porites_Sthermophi
C_Zam03_28_10m_C27
C_Baja04_124_C1_Ppanamensis
C_Cur11_33_C7
I_FN561562_Sym_cladeI
I_FN561559_Sym_cladeI
G_AJ291537_Sym_clade_G
G_AJ291538_Sym_clade_G
G_AJ291539_Sym_clade_G
G_OI_C_orient_1
G_Be02_183
D_AJ308899.1_PSP1_05_sponge
D_JQ247049_1_marginopora_D1_1
D_rt401_MH301113
D_Peng12_03_D15_S_boreaum
D_Pal09_63_D1_S_glynnii
D_Tha09_66_D8_S_eurythalpos
D_CCMP_2556_Strenchii
E_RCC1521_tcla
E_CCMP_3410_rt383_Svoratum
E_CCMP_421_tcl
A_MED15_Anemonia_viridis48t_te
A_BI16_Cg_temporate_A
A_Med15_Paranemonia_cinerea27_
A_NZ_anthopleura_temporate_A
A_CCMP_2461_S_pilosum_rt185
A_CCMP_2548_Snatans_Pacific
A_CCMP_2456_typeA4_rt379
A_CCMP_2457_Stridacnidorum
A_CCMP_2465_rt292_A3_Stridacni
A_CCMP_2469_Snecroappetens_rt8
A_CCMP_2464_rt61Smicroadriatic
0.1
HG529980_Ansanella_granifera_A
EF052683_Baldinia_anauniensis
AY571374_Borghiella_tenuissima
EU126801_Borghiella_dodgei
LC068841_Leiocephalium_pseudos
LN898222_Yihiella_yeosuensis_Y
AY036081_Polarellaglacialis_FL
AY628430_Biecheleria_halophila
AF260402_Biecheleria_pseudopal
FJ024705_Biecheleria_cincta
DQ195364_Pelagodinium_beii_DVf
KP342301_Pelagodinium_sp_EP_20
EF205014_Protodinium_simplex_m
GQ422124_Pelagodinium_beii_Sai
GQ422124_Pelagodinium_beii
KM603188_Biecheleriopsis_adria
EU857537_Biecheleriopsis_adria
AF318226_Protodinium_coriii
AF060901_Protodinium_simplex_C
Breviolum
Fugacium
Cladocopium
Gerakladium
Durusdinium
Effrenium
Symbiodinium

Figure S1-2. Effects of heat stress on symbiotic *E. diaphana*

Measurement of each heat stress index after incubation at 33°C for 24 h. Values are mean ± SEM. A. The number of symbiont cells per host protein (n = 5 per treatment, *t-test*); B. Maximum quantum yield of photosystem II (n = 4 per treatment, paired *t-test*); C. Photosynthesis rate (n = 4 per treatment, paired *t-test*); D. Respiration rate (n = 4 per treatment, paired *t-test*).

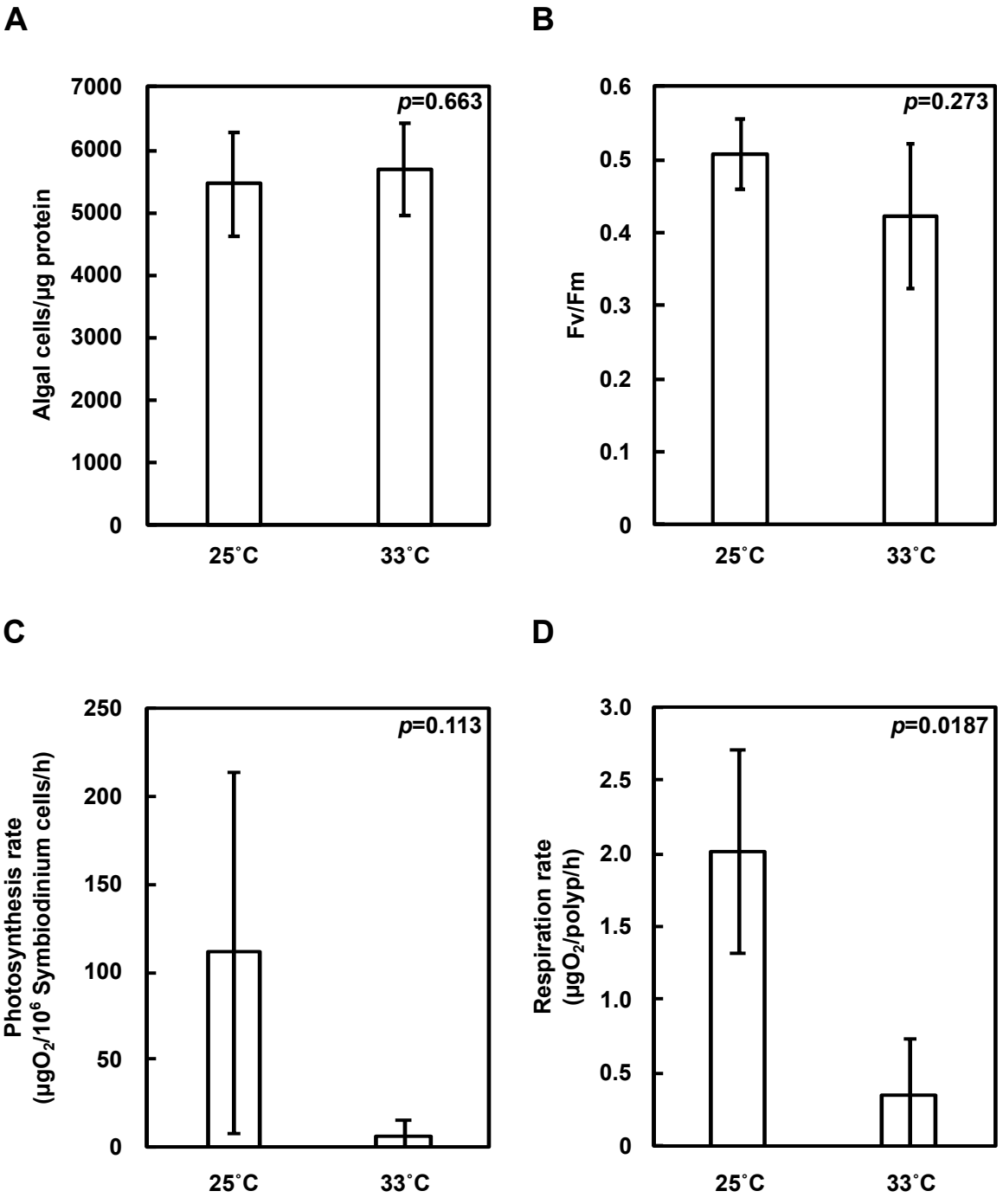
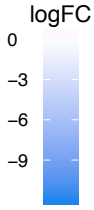


Figure S1-3. Differences of expression levels of *NPC2* genes in different conditions

Frames in bold indicate DEGs between two conditions.



Sym-Heat / Sym-Norm	Apo-Heat / Apo-Norm	Apo-Heat / Sym-Norm	Apo-Norm / Sym-Norm		
-0.65	-1.08	-2.23	-1.17	NPC2A	AIPGENE3167
-0.83	0.39	-2.61	-3.01	NPC2F	AIPGENE3074
-1.61	-0.03	-10.88	-10.87	NPC2B	AIPGENE22527
-1.23	-1.20	-11.97	-10.79	NPC2D	AIPGENE22473
-0.73	-0.85	-4.92	-4.09	NPC2C	AIPGENE22539
-1.13	-0.44	-2.03	-1.61	NPC2E	AIPGENE20032

Figure S1-4. Phylogenetic tree of NPC2-type sterol transporter proteins.

A maximum likelihood tree of NPC2 proteins with species names and GenBank IDs is shown. Support values using an SH-like approximate likelihood ratio test (left) and ultrafast bootstrap approximation (right) are shown on each branch. For *E. diaphana*, the 'AIPGENE' gene IDs and gene names are used according to a previous study (Baumgarten *et al.* 2015). Arrowheads indicate genes for which expression was differentially expressed between different temperatures in the symbiotic individuals and double circles indicates DEGs between different symbiotic states under normal temperature, in this study.

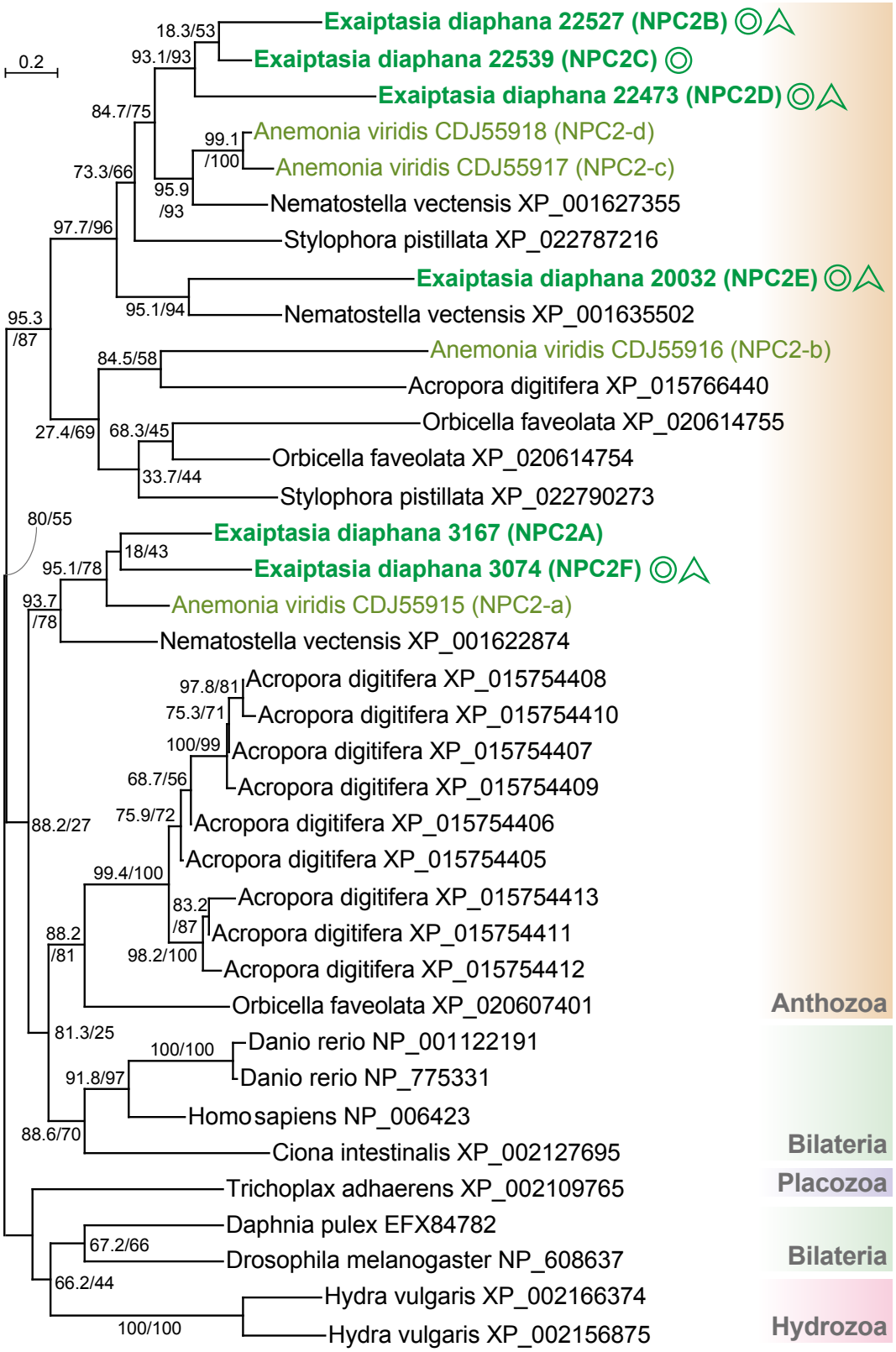

0.2
Exaiptasia diaphana 22527 (NPC2B)
Exaiptasia diaphana 22539 (NPC2C)
Exaiptasia diaphana 22473 (NPC2D)
Anemonia viridis CDJ55918 (NPC2-d)
Anemonia viridis CDJ55917 (NPC2-c)
Nematostella vectensis XP_001627355
Stylophora pistillata XP_022787216
Exaiptasia diaphana 20032 (NPC2E)
Nematostella vectensis XP_001635502
Anemonia viridis CDJ55916 (NPC2-b)
Acropora digitifera XP_015766440
Orbicella faveolata XP_020614755
Orbicella faveolata XP_020614754
Stylophora pistillata XP_022790273
Exaiptasia diaphana 3167 (NPC2A)
Exaiptasia diaphana 3074 (NPC2F)
Anemonia viridis CDJ55915 (NPC2-a)
Nematostella vectensis XP_001622874
Acropora digitifera XP_015754408
Acropora digitifera XP_015754410
Acropora digitifera XP_015754407
Acropora digitifera XP_015754409
Acropora digitifera XP_015754406
Acropora digitifera XP_015754405
Acropora digitifera XP_015754413
Acropora digitifera XP_015754411
Acropora digitifera XP_015754412
Orbicella faveolata XP_020607401
Danio rerio NP_001122191
Danio rerio NP_775331
Homo sapiens NP_006423
Ciona intestinalis XP_002127695
Trichoplax adhaerens XP_002109765
Daphnia pulex EFX84782
Drosophila melanogaster NP_608637
Hydra vulgaris XP_002166374
Hydra vulgaris XP_002156875
Anthozoa
Bilateria
Placozoa
Bilateria
Hydrozoa

Figure S1-5. Presence–absence matrix and expression level in HR-DEGs shared by symbiotic states

HR-DEGs associated with enriched GO terms are shown as in Figure 1-3C.

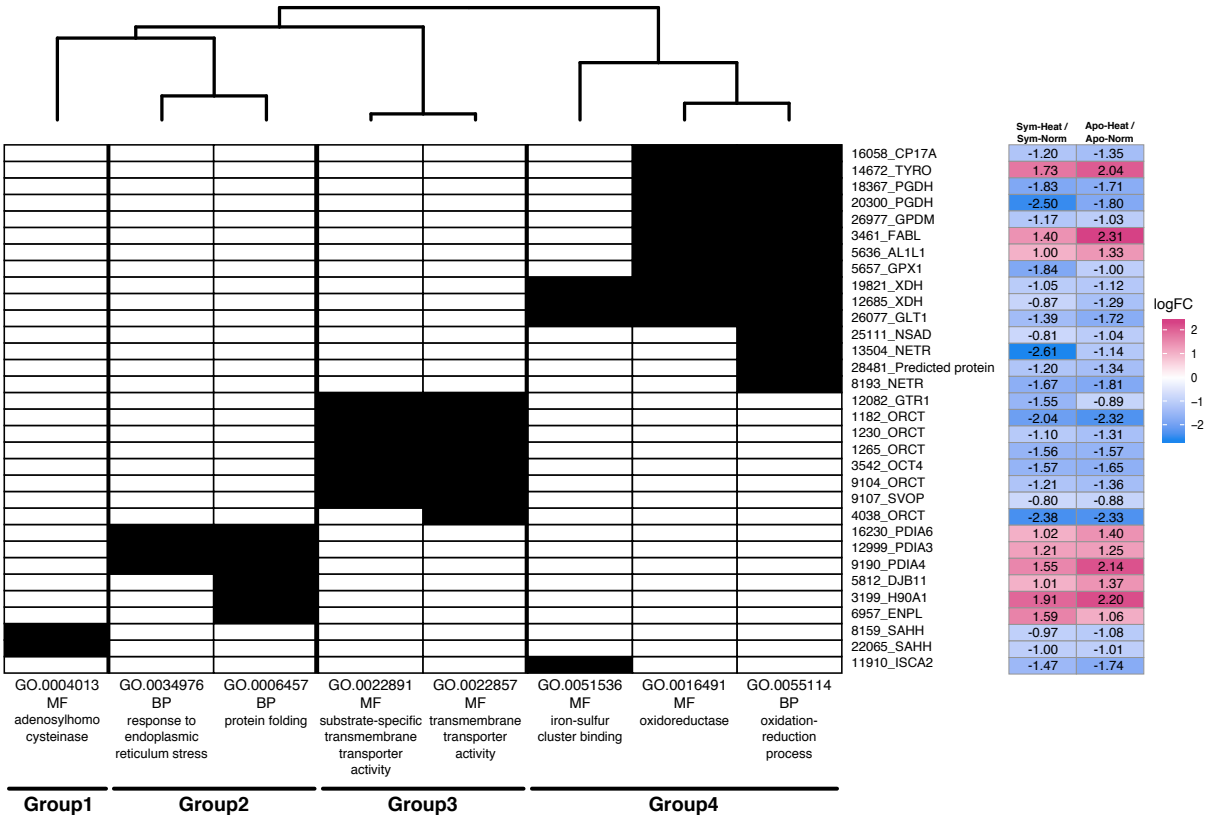
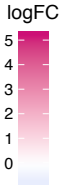


Figure S1-6. Differences of expression levels of HSP genes in different conditions

Frames in bold indicate DEGs between two conditions.



Sym-Heat / Sym-Norm	Apo-Heat / Apo-Norm			
0.50	**1.05**	AHSA1	AIPGENE9489	Activator of 90 kDa heat shock protein ATPase homolog 1
0.69	**1.17**	CH10	AIPGENE15266	10 kDa heat shock protein, mitochondrial
0.04	-0.25	CH60	AIPGENE15267	60 kDa heat shock protein, mitochondrial
-0.03	0.38	CHSP1	AIPGENE12745	Calcium-regulated heat stable protein 1
1.91	**2.20**	H90A1	AIPGENE3199	Heat shock protein HSP 90-alpha 1
0.70	0.42	HAP28	AIPGENE12829	28 kDa heat- and acid-stable phosphoprotein
-0.49	0.14	HS12A	AIPGENE6574	Heat shock 70 kDa protein 12A
-0.71	-0.30	HS12B	AIPGENE13832	Heat shock 70 kDa protein 12B
2.96	1.18	HS71A	AIPGENE18790	Heat shock 70 kDa protein 1A
0.35	0.48	HSF	AIPGENE14111	Heat shock factor protein
-0.40	-0.45	HSF1	AIPGENE11173	Heat shock factor protein 1
-0.01	-0.57	HSF2B	AIPGENE13751	Heat shock factor 2-binding protein
0.94	1.76	HSF4	AIPGENE8965	Heat shock factor protein 4
-0.06	0.14	HSP13	AIPGENE22153	Heat shock 70 kDa protein 13
1.61	0.76	HSP70	AIPGENE8387	Heat shock 70 kDa protein
5.21	1.13	HSP7C	AIPGENE21738	Heat shock cognate 71 kDa protein
0.91	0.52	HSP7C	AIPGENE5814	Heat shock cognate 71 kDa protein
0.31	0.95	HSP7E	AIPGENE21444	Heat shock 70 kDa protein 14
1.02	0.67	HSP97	AIPGENE27608	97 kDa heat shock protein
0.23	0.69	TRAP1	AIPGENE15468	Heat shock protein 75 kDa, mitochondrial

Figure S1-7. Presence–absence matrix and expression level in HR-DEGs unique to the symbiotic individuals

HR-DEGs associated with enriched GO terms are shown as in Figure 1-3C.

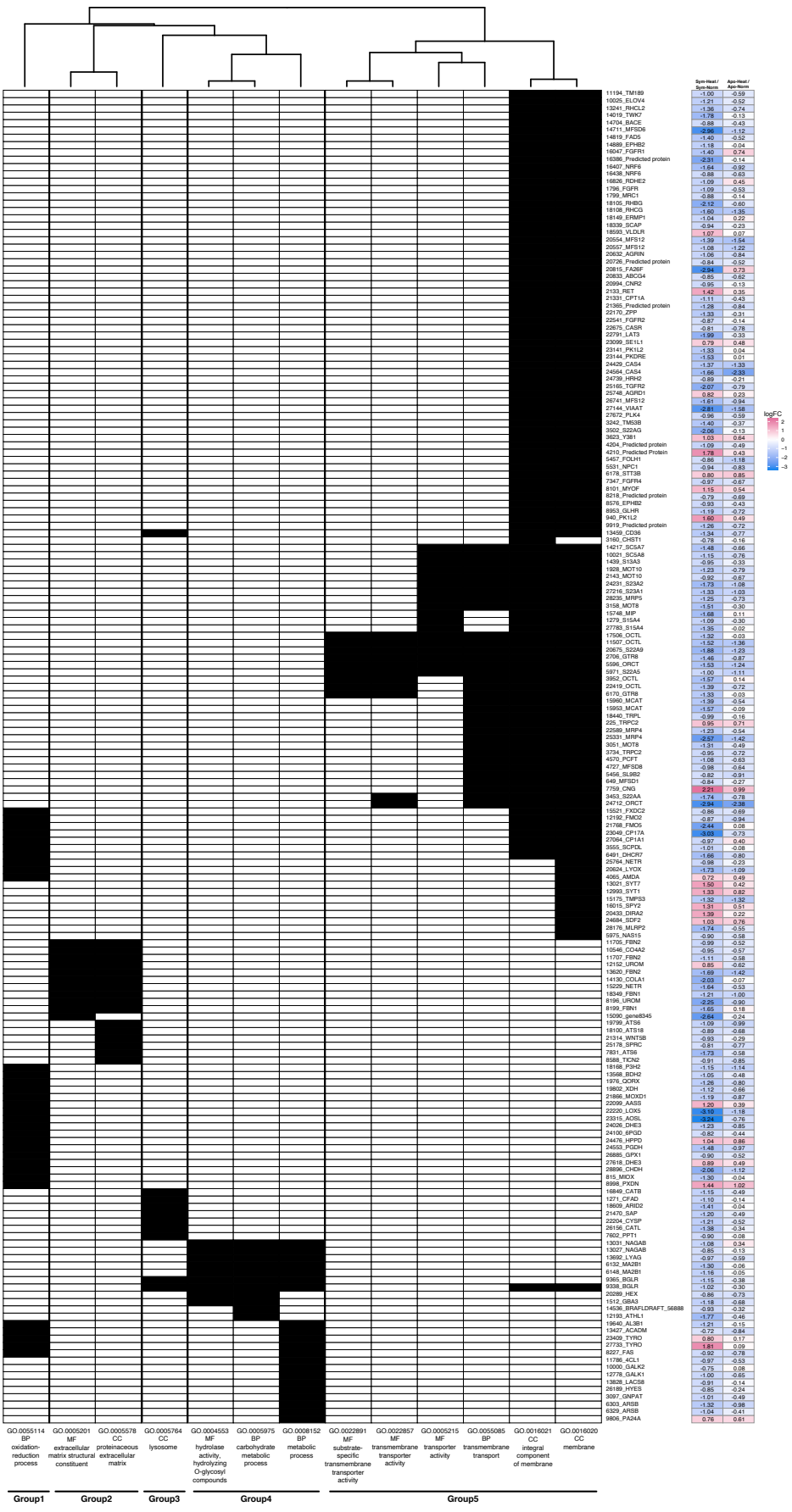
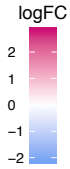

Sym-Heat / Sym-Norm
Apo-Heat / Apo-Norm
logFC
2
1
0
−1
−2
−3
GO.0055114 BP oxidation-reduction process
GO.0005201 MF extracellular matrix structural constituent
GO.0005578 CC proteinaceous extracellular matrix
GO.0005764 CC lysosome
GO.0004553 MF hydrolase activity, hydrolyzing O-glycosyl compounds
GO.0005975 BP carbohydrate metabolic process
GO.0008152 BP metabolic process
GO.0022891 MF substrate-specific transmembrane transporter activity
GO.0022857 MF transmembrane transporter activity
GO.0005215 MF transporter activity
GO.0055085 BP transmembrane transport
GO.0016021 CC integral component of membrane
GO.0016020 CC membrane
Group1
Group2
Group3
Group4
Group5

Figure S1-8. Differences of expression levels of Rab family genes in different conditions

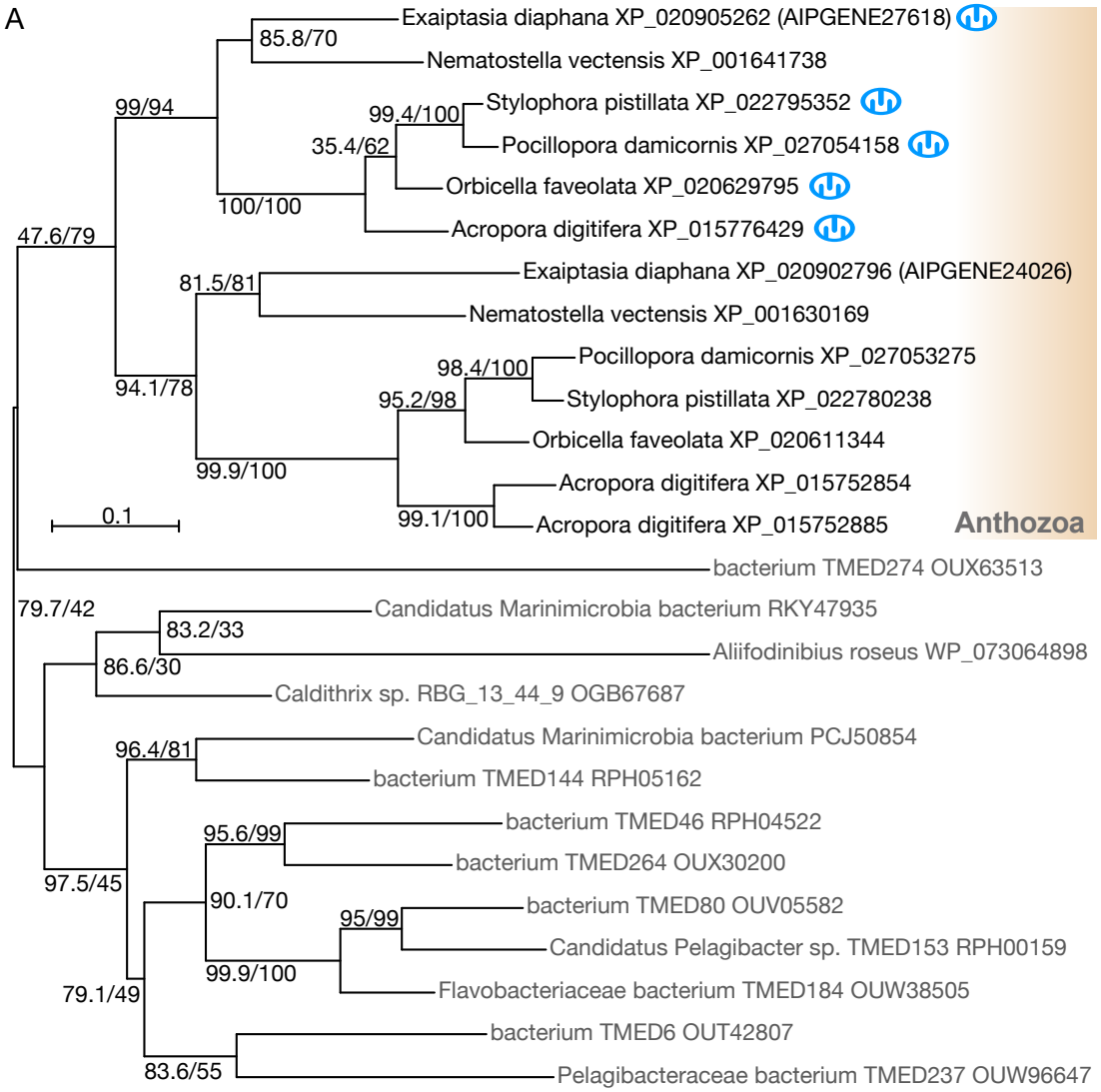
Frames in bold indicate DEGs between two conditions.



Sym-Heat / Sym-Norm	Apo-Heat / Apo-Norm	Apo-Heat / Sym-Norm	Apo-Norm / Sym-Norm			
0.60	0.46	0.96	0.48	GORAB	AIPGENE2803	RAB6-interacting golgin
-0.15	0.48	-0.41	-0.91	RAB13	AIPGENE12806	Ras-related protein Rab-13
-0.34	0.33	**-1.04**	**-1.38**	RAB13	AIPGENE12823	Ras-related protein Rab-13
0.49	0.17	0.07	-0.11	RAB13	AIPGENE18849	Ras-related protein Rab-13
0.17	-0.14	-0.59	-0.46	RAB13	AIPGENE19750	Ras-related protein Rab-13
0.21	0.32	-0.06	-0.40	RAB14	AIPGENE18421	Ras-related protein Rab-14
-0.32	0.60	**-0.81**	**-1.42**	RAB18	AIPGENE11215	Ras-related protein Rab-18
0.21	0.24	-0.23	-0.49	RAB1A	AIPGENE2448	Ras-related protein Rab-1A
0.48	0.17	0.43	0.25	RAB1B	AIPGENE16067	Ras-related protein Rab-1B
0.24	0.28	-0.18	-0.47	RAB2	AIPGENE3537	Ras-related protein Rab-2
-0.09	0.01	-0.63	-0.66	RAB20	AIPGENE8885	Ras-related protein Rab-20
-0.13	0.27	-0.48	-0.77	RAB21	AIPGENE11177	Ras-related protein Rab-21
0.00	0.32	-0.22	-0.55	RAB21	AIPGENE11551	Ras-related protein Rab-21
-1.61	0.94	-0.95	-1.90	RAB23	AIPGENE28809	Ras-related protein Rab-23
-0.17	-1.29	-0.59	0.68	RAB23	AIPGENE28829	Ras-related protein Rab-23
0.14	0.19	-0.47	-0.68	RAB24	AIPGENE24242	Ras-related protein Rab-24
-0.63	-0.07	**-1.36**	**-1.31**	RAB24	AIPGENE24297	Ras-related protein Rab-24
-0.45	0.45	**-1.22**	**-1.68**	RAB24	AIPGENE24304	Ras-related protein Rab-24
-0.55	0.18	**-0.80**	**-1.00**	RAB24	AIPGENE9333	Ras-related protein Rab-24
0.64	0.50	0.21	-0.31	RAB26	AIPGENE17105	Ras-related protein Rab-26
-1.09	-0.40	**-2.10**	**-1.71**	RAB26	AIPGENE28373	Ras-related protein Rab-26
-0.14	-0.30	-0.03	0.26	RAB28	AIPGENE23890	Ras-related protein Rab-28
0.01	0.26	-0.08	-0.36	RAB3	AIPGENE25566	Ras-related protein Rab-3
0.04	-0.49	-0.22	0.26	RAB30	AIPGENE201	Ras-related protein Rab-30
-0.13	-0.16	-0.53	-0.38	RAB30	AIPGENE26838	Ras-related protein Rab-30
-1.08	-0.33	**-1.54**	**-1.23**	RAB32	AIPGENE20314	Ras-related protein Rab-32
0.20	0.07	-0.36	-0.44	RAB33	AIPGENE207	Putative Ras-related protein Rab-33
-0.04	0.19	-0.55	-0.76	RAB35	AIPGENE16958	Ras-related protein Rab-35
-0.09	-0.04	-0.63	-0.60	RAB36	AIPGENE27079	Ras-related protein Rab-36
0.09	-0.12	-0.40	-0.30	RAB3A	AIPGENE7581	Ras-related protein Rab-3A
-0.91	-0.83	-0.61	0.21	RAB43	AIPGENE13082	Ras-related protein Rab-43
0.06	0.26	0.04	-0.24	RAB4B	AIPGENE12050	Ras-related protein Rab-4B
-0.24	0.14	**-0.69**	**-0.85**	RAB5B	AIPGENE10564	Ras-related protein Rab-5B
0.08	0.34	-0.37	-0.72	RAB6	AIPGENE13835	Ras-related protein Rab6
-0.36	-0.20	-0.41	-0.22	RAB7A	AIPGENE5165	Ras-related protein Rab-7a
-0.92	-0.42	**-1.25**	**-0.84**	RAB7L	AIPGENE10740	Ras-related protein Rab-7L1
-0.17	-0.14	-0.59	-0.46	RAB7L	AIPGENE12601	Ras-related protein Rab-7L1
2.79	-1.36	1.46	2.81	RAB8A	AIPGENE13722	Ras-related protein Rab-8A
0.00	0.12	-0.21	-0.34	RAB9A	AIPGENE27699	Ras-related protein Rab-9A
0.01	0.10	0.48	0.36	RABE1	AIPGENE9062	Rab GTPase-binding effector protein 1
-0.30	0.08	-0.72	-0.81	RABEK	AIPGENE24303	Rab9 effector protein with kelch motifs
0.30	0.16	0.80	0.62	RABL6	AIPGENE25499	Rab-like protein 6
0.02	-0.30	0.19	0.47	RABX5	AIPGENE25911	Rab5 GDP/GTP exchange factor

Figure S1-9. Sequence analyses of cnidarian glutamate dehydrogenase

A maximum likelihood tree of glutamate dehydrogenase DHE3 proteins. Support values using an SH-like approximate likelihood ratio test (left) and ultrafast bootstrap approximation (right) are shown on each branch. 'M' in oval (blue) indicates the presence of mitochondrial targeting peptide predicted by iPSORT. B. Amino acid alignment of cnidarian glutamate dehydrogenases.



B

Predicted mitochondrial targeting peptide

```
XP_020905262 (AIPGENE27618)-----MLR--VXGRG-TFSLGR- VAIRC--LS -----TTT--LNAAPSGT STKISEPKYK SIDDFMGYVE KKNPGEVELH
XP_022795352 Stylophora     MFLRSVARRFCIQRGSSLLLGR- VAVRC--AS -----GVY--RTAMMENG EMSDGNMRFK GVDDFMNHVV KKNPGEHEFH
XP_027054158 Pocillopora    MFLRSVARGFSAQRGLSLLLGR- VAVRC--AS -----GIYPNPTASVENG EMSDGNMKFK CVDDFMNHVV KRNPGEHEFH
XP_020629795 Orbicella      MFLRTIGSRFSIQRG-SR----- LLRRV--PV -RCITGVQ-QSAIIENGD TLVEGNMKFK SLDDFMSHVV TKNPGEHEFH
XP_015776429 Acropora       M---SFV---------------- KV-I-RIVT -QEVSRKNKWMLNGLSSR SGPSNQMRYK SIDDFMNHVE QKNPGEHEFH
XP_001641738 Nematostella   ----------------------- --------- -----------------M WNQHLEGQYK SPEDFLETVR RKNPGEAEFL

XP_020902796 (AIPGENE24026)----------------------- --------- ------------------ ---MDCCGCK SPDDFLQLVK RNNPGEDEFH
XP_022780238 Stylophora     ----------------------- --------- ------------------ ----MAGKYA SPEEFLKYVK TRSPGEDEFH
XP_027053275 Pocillopora    ----------------------- --------- ------------------ ----MEGKYA SPEEFLEYVK MRSPGEDEFH
XP_020611344 Orbicella      ----------------------- --------- ------------------ ----MSSKYS SPEEFLDYVK RKNPGEHEFH
XP_015752854 Acropora       ----------------------- --------- ------------------ -MSETESKYK SPEAFVNYVV RKNPGEKEFH
XP_001630169 Nematostella   ----------------------- --------- ------------------ ----MGCRFN SPEEFLKFVK AKDPQESEFH
```

N-terminal extension | Mature protein

Table S1-1. Enriched GO terms for the host DEGs shared by different symbiotic states

E. diaphana geneID	Gene name	Discription	UniprotID	*C. intestinalis* geneID	GOID
AIPGENE3542	4-Oct	Organic cation/carnitine transporter 4	Q9LHQ6	ENSCINP00000001843.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE21249	ABCG2	ATP-binding cassette sub-family G member 2	Q80W57	ENSCINP00000008035.3	GO:0000166 GO:0005524 GO:0016020 GO:0016021 GO:0016887
AIPGENE25480	ACES	Acetylcholinesterase	O62763	ENSCINP00000031848.1	
AIPGENE9142	ACTC	Actin, cytoplasmic	P12716	ENSCINP00000001538.3	GO:0000166 GO:0005524
AIPGENE4946	AGRIN	Agrin	A2ASQ1	ENSCINP00000024219.2	
AIPGENE5636	AL1L1	Cytosolic 10-formyltetrahydrofolate dehydrogenase	Q6GNL7	ENSCINP00000016529.3	GO:0003824 GO:0005737 GO:0006730 GO:0008152 GO:0009058 GO:0009258 GO:0016155 GO:0016491 GO:0016620 GO:0016742 GO:0019145 GO:0047105 GO:0055114
AIPGENE17420	AMT1	Putative ammonium transporter 1	P54145	ENSCINP00000012012.3	GO:0005887 GO:0006810 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488
AIPGENE19941	ARLZ	Probable argininosuccinate lyase	P50514	ENSCINP00000013034.3	GO:0003824 GO:0004056 GO:0005829 GO:0006526 GO:0042450 GO:0051262
AIPGENE8629	ARMET	Mesencephalic astrocyte-derived neurotrophic factor homolog	B4QX46	ENSCINP00000031474.1	
AIPGENE2483	ATHL1	Acid trehalase-like protein 1	A0JMP0	ENSCINP00000005930.3	GO:0003824 GO:0005975
AIPGENE15022	BETT	High-affinity choline transport protein	P0ABC9		
AIPGENE22194	CALL3	Calmodulin-like protein 3	Q5U206	ENSCINP00000017574.3	
AIPGENE9256	CALR	Calreticulin	P29413	ENSCINP00000016510.3	
AIPGENE25880	CALUA	Calumenin-A	B5X186	ENSCINP00000025165.2	GO:0005509
AIPGENE308	CALUA	Calumenin-A	Q6IQP3	ENSCINP00000007993.2	GO:0005509
AIPGENE9938	CALUB	Calumenin-B	B5X4E0	ENSCINP00000025165.2	GO:0005509
AIPGENE13229	CBPA2	Carboxypeptidase A2	Q504N0	ENSCINP00000011346.3	GO:0004180 GO:0004181 GO:0005615 GO:0006508 GO:0008270
AIPGENE13576	CBPA4	Carboxypeptidase A4	Q9UI42	ENSCINP00000011346.3	GO:0004180 GO:0004181 GO:0005615 GO:0006508 GO:0008270
AIPGENE18534	CO1A2	Collagen alpha-2(I) chain	P02466	ENSCINP00000003025.3	GO:0005201 GO:0005578 GO:0005581
AIPGENE25170	CO1A2	Collagen alpha-2(I) chain	O42350	ENSCINP00000014311.3	GO:0005201 GO:0005581
AIPGENE25192	CO1A2	Collagen alpha-2(I) chain	P08123	ENSCINP00000017221.3	

AIPGENE8376	CO1A2	Collagen alpha-2(I) chain	P02466	ENSCINP00000017221.3	
AIPGENE5608	CO4A1	Collagen alpha-1(IV) chain	P17139	ENSCINP00000016017.3	
AIPGENE12347	CO6A6	Collagen alpha-6(VI) chain	A6NMZ7	ENSCINP00000031657.1	
AIPGENE25189	COA	Collagen alpha chain (Fragment)	B8V7R6	ENSCINP00000017221.3	
AIPGENE1554	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	P30437	ENSCINP00000032679.1	
AIPGENE16058	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	O73853	ENSCINP00000018595.3	GO:0004497 GO:0005506 GO:0008395 GO:0016020 GO:0016021 GO:0016491 GO:0016705 GO:0020037 GO:0046872 GO:0055114
AIPGENE15958	CP341	Cytochrome P450 3A41	Q9JMA7	ENSCINP00000036138.1	
AIPGENE25863	CRA1B	Collagen alpha-1(XXVII) chain B	A0MSJ1	ENSCINP00000016017.3	
AIPGENE25893	CRA1B	Collagen alpha-1(XXVII) chain B	A0MSJ1	ENSCINP00000017221.3	
AIPGENE6906	CREL2	Cysteine-rich with EGF-like domain protein 2	Q7SXF6	ENSCINP00000024648.2	GO:0005509 GO:0016020 GO:0016021
AIPGENE7700	CTP5B	Contactin-associated protein like 5-2	Q0V8T5	ENSCINP00000010501.3	
AIPGENE2858	DDX5	Probable ATP-dependent RNA helicase DDX5	Q61656	ENSCINP00000008570.3	GO:0000166 GO:0003676 GO:0004004 GO:0004386 GO:0005524 GO:0010501 GO:0016787 GO:0045893
AIPGENE26806	DEFM	Peptide deformylase, mitochondrial	Q9HBH1		
AIPGENE11883	DEGS1	Sphingolipid delta(4)-desaturase DES1	Q3ZBY7	ENSCINP00000030732.1	
AIPGENE5812	DJB11	DnaJ homolog subfamily B member 11	Q6TUG0	ENSCINP00000032640.1	GO:0006457 GO:0051082
AIPGENE25828	DNJC3	DnaJ homolog subfamily C member 3	Q27968	ENSCINP00000019020.3	
AIPGENE18998	DPEP1	Dipeptidase 1	P16444	ENSCINP00000028153.2	
AIPGENE6957	ENPL	Endoplasmin	Q66HD0	ENSCINP00000015093.3	GO:0005524 GO:0006457 GO:0006950 GO:0051082
AIPGENE2929	FA8	Coagulation factor VIII	P12263	ENSCINP00000006532.3	GO:0005509
AIPGENE3461	FABL	Enoyl-[acyl-carrier-protein] reductase [NADPH] FabL	P71079	ENSCINP00000018876.3	GO:0016491 GO:0055114
AIPGENE4474	FGFR	Fibroblast growth factor receptor	Q26614	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE17494	FGFR4	Fibroblast growth factor receptor 4	Q498D6	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108

AIPGENE20955	FKBP2	Peptidyl-prolyl cis-trans isomerase FKBP2	P45878	ENSCINP00000030674.1	
AIPGENE6958	FP	Fibronectin type III domain-containing protein	B8V1W9		
AIPGENE22443	FRIS	Soma ferritin	P42577	ENSCINP00000035294.1	
AIPGENE25324	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000001814.3	GO:0005576
AIPGENE8921	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9H4G4	ENSCINP00000010145.3	GO:0005576
AIPGENE26077	GLT1	Putative glutamate synthase [NADPH]	Q9C102	ENSCINP00000007792.3	GO:0003824 GO:0005737 GO:0006212 GO:0016491 GO:0016627 GO:0017113 GO:0051536 GO:0055114
AIPGENE22837	GMDS	GDP-mannose 4,6 dehydratase	Q8K3X3	ENSCINP00000010294.3	GO:0008446 GO:0019673
AIPGENE26977	GPDM	Glycerol-3-phosphate dehydrogenase, mitochondrial	A6QLU1	ENSCINP00000002723.3	GO:0004368 GO:0005509 GO:0006072 GO:0009331 GO:0016491 GO:0052590 GO:0052591 GO:0055114
AIPGENE5657	GPX1	Glutathione peroxidase	Q00277	ENSCINP00000035821.1	GO:0004601 GO:0004602 GO:0006979 GO:0016491 GO:0055114 GO:0098869
AIPGENE12496	GRP78	78 kDa glucose-regulated protein	Q90593	ENSCINP00000019494.3	GO:0000166 GO:0005524
AIPGENE6332	GSTM1	Glutathione S-transferase Mu 1	Q9N0V4	ENSCINP00000010811.1	GO:0004364 GO:0008152
AIPGENE12082	GTR1	Solute carrier family 2, facilitated glucose transporter member 1	P20303	ENSCINP00000003472.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE3199	H90A1	Heat shock protein HSP 90-alpha 1	Q90474	ENSCINP00000023416.2	GO:0005524 GO:0006457 GO:0006950 GO:0051082
AIPGENE15071	HE	Hatching enzyme	P22757	ENSCINP00000017258.3	
AIPGENE4831	HELT	Hairy and enhancer of split-related protein HELT	A6NFD8	ENSCINP00000013234.3	GO:0003677 GO:0005634 GO:0006351 GO:0006355 GO:0046983
AIPGENE157	HMGB3	High mobility group protein B3	Q32L31	ENSCINP00000007915.3	
AIPGENE27402	HMGX3	HMG domain-containing protein 3	Q12766	ENSCINP00000030465.1	
AIPGENE22432	HSP31	Glyoxalase 3	Q5AF03		
AIPGENE651	HYOU1	Hypoxia up-regulated protein 1	Q7ZUW2	ENSCINP00000031737.1	GO:0000166 GO:0005524
AIPGENE3465	IF44L	Interferon-induced protein 44-like	Q53G44	ENSCINP00000019144.3	
AIPGENE11910	ISCA2	Iron-sulfur cluster assembly 2 homolog, mitochondrial	Q86U28	ENSCINP00000036261.1	GO:0005198 GO:0006790 GO:0008198 GO:0016226 GO:0051536 GO:0051537 GO:0051539 GO:0097428

AIPGENE15815	ISCU	Iron-sulfur cluster assembly enzyme ISCU, mitochondrial	Q9D7P6	ENSCINP00000031409.1	
AIPGENE22835	JMJCD	JmjC domain-containing protein D	Q55DF5	ENSCINP00000025973.2	
AIPGENE7475	JMJD6	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6	Q5ZMK5	ENSCINP00000015483.3	
AIPGENE7692	KCRS	Creatine kinase S-type, mitochondrial	P11009	ENSCINP00000017993.3	
AIPGENE26035	KDM8	Lysine-specific demethylase 8	B2GUS6	ENSCINP00000010139.3	GO:0005509
AIPGENE6242	KDM8	Lysine-specific demethylase 8	B2GUS6	ENSCINP00000010188.3	GO:0005509
AIPGENE7224	KLKB1	Plasma kallikrein	P26262	ENSCINP00000025876.2	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE21474	LICH	Putative lysosomal acid lipase/cholesteryl ester hydrolase	J3SDX8	ENSCINP00000032128.1	GO:0016042 GO:0016787 GO:0016788
AIPGENE3723	LMAN1	Protein ERGIC-53	Q9TU32	ENSCINP00000005331.3	GO:0016020 GO:0016021
AIPGENE13248	LONF3	LON peptidase N-terminal domain and RING finger protein 3	Q496Y0	ENSCINP00000026615.2	
AIPGENE4997	MANF	Mesencephalic astrocyte-derived neurotrophic factor	P55145	ENSCINP00000031474.1	
AIPGENE9298	MEP1B	Meprin A subunit beta	Q16820	ENSCINP00000017081.3	GO:0004222 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016020 GO:0016787 GO:0046872
AIPGENE11428	METB	Cystathionine gamma-synthase	P56069	ENSCINP00000025947.2	GO:0003824 GO:0030170
AIPGENE14224	MIB	E3 ubiquitin-protein ligase mind-bomb	Q9VUX2	ENSCINP00000022163.2	
AIPGENE14222	MIB1	E3 ubiquitin-protein ligase MIB1	Q86YT6	ENSCINP00000001354.3	GO:0004842 GO:0005737 GO:0006897 GO:0007219 GO:0008270 GO:0016567 GO:0046872
AIPGENE14267	MIB1	E3 ubiquitin-protein ligase MIB1	Q86YT6	ENSCINP00000001354.3	GO:0004842 GO:0005737 GO:0006897 GO:0007219 GO:0008270 GO:0016567 GO:0046872
AIPGENE14234	MIB2	E3 ubiquitin-protein ligase MIB2	Q5ZIJ9	ENSCINP00000001354.3	GO:0004842 GO:0005737 GO:0006897 GO:0007219 GO:0008270 GO:0016567 GO:0046872
AIPGENE22864	MLP	Mucin-like protein (Fragment)	B3EWY9	ENSCINP00000001942.3	GO:0005509
AIPGENE9728	MMEL1	Membrane metallo-endopeptidase-like 1	Q495T6	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE2999	MOT10	Monocarboxylate transporter 10	Q91Y77	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE5882	MTPN	Myotrophin	Q6P1S6	ENSCINP00000009430.3	GO:0005737

					GO:0008093 GO:0030507 GO:0072661
AIPGENE13355	MZB1	Marginal zone B- and B1-cell-specific protein	Q8WU39	ENSCINP00000031626.1	
AIPGENE13504	NETR	Neurotrypsin	O08762	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE8193	NETR	Neurotrypsin	G3V801	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE22061	NFS1	Cysteine desulfurase, mitochondrial	Q5RDE7	ENSCINP00000023238.2	
AIPGENE4218	NHL1	RING finger protein nhl-1	Q03601	ENSCINP00000001697.3	GO:0005622 GO:0008270 GO:0046872
AIPGENE27652	NOTC3	Neurogenic locus notch homolog protein 3	Q9R172	ENSCINP00000000877.3	GO:0005509
AIPGENE5533	NPC1	Niemann-Pick C1 protein	O35604	ENSCINP00000008700.3	GO:0005319 GO:0006869 GO:0016020 GO:0016021
AIPGENE16104	NPHP3	Nephrocystin-3	Q7Z494	ENSCINP00000034474.1	
AIPGENE25111	NSAD	2-hydroxychromene-2-carboxylate isomerase	Q9X9Q7	ENSCINP00000030897.1	GO:0015035 GO:0055114
AIPGENE1182	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000015651.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE1230	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000020124.3	GO:0005887 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE1265	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000007897.3	GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE4038	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000035947.1	GO:0005452 GO:0005887 GO:0015347 GO:0015698 GO:0016020 GO:0016021 GO:0022857 GO:0043252 GO:0055085
AIPGENE9104	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000007897.3	GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE14447	PAG15	Group XV phospholipase A2	Q8NCC3	ENSCINP00000006478.3	GO:0006629 GO:0008374
AIPGENE27122	PCY2	Ethanolamine-phosphate cytidylyltransferase	Q99447	ENSCINP00000006470.3	GO:0003824 GO:0009058
AIPGENE12999	PDIA3	Protein disulfide-isomerase A3	Q5RDG4	ENSCINP00000005773.4	GO:0003756 GO:0005783 GO:0006457 GO:0009986 GO:0016853 GO:0034976 GO:0045454
AIPGENE9190	PDIA4	Protein disulfide-	P13667	ENSCINP00000013204.3	GO:0003756

		isomerase A4			GO:0005783 GO:0006457 GO:0016853 GO:0034976 GO:0045454
AIPGENE16230	PDIA6	Protein disulfide-isomerase A6	Q63081	ENSCINP00000008588.3	GO:0003756 GO:0005783 GO:0006457 GO:0016853 GO:0034976 GO:0043277 GO:0045454
AIPGENE18367	PGDH	15-hydroxyprostaglandin dehydrogenase [NAD(+)]	Q3T0C2	ENSCINP00000003988.3	GO:0016491 GO:0055114
AIPGENE20300	PGDH	15-hydroxyprostaglandin dehydrogenase [NAD(+)]	P70684	ENSCINP00000011563.3	GO:0016491 GO:0055114
AIPGENE15654	PGFS	Prostamide/prostaglandin F synthase	B5X9L9	ENSCINP00000025320.2	
AIPGENE4566	PIF1	ATP-dependent DNA helicase PIF1	Q5AXT5	ENSCINP00000005776.3	GO:0000002 GO:0000166 GO:0000723 GO:0003677 GO:0003678 GO:0004386 GO:0005524 GO:0005634 GO:0005657 GO:0005739 GO:0006260 GO:0006281 GO:0006310 GO:0006974 GO:0010521 GO:0016787 GO:0032211 GO:0032508 GO:0043141 GO:0044806 GO:0051974
AIPGENE7117	POL2	Retrovirus-related Pol polyprotein from transposon 297	P20825	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE20500	POL3	Retrovirus-related Pol polyprotein from transposon 17.6	P04323	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE18183	PPD	Phosphonopyruvate decarboxylase	Q54271		
AIPGENE18187	PPD	Phosphonopyruvate decarboxylase	Q54271		
AIPGENE8936	PRPX	Pathogen-related protein	P16273		
AIPGENE5760	PTPRQ	Phosphatidylinositol phosphatase PTPRQ	O88488	ENSCINP00000015322.3	
AIPGENE25618	RAVR2	Ribonucleoprotein PTB-binding 2	Q9HCJ3	ENSCINP00000005147.3	GO:0000166 GO:0000398 GO:0003676 GO:0005634
AIPGENE27196	S23A2	Solute carrier family 23 member 2	Q9EPR4	ENSCINP00000007695.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE27243	S23A2	Solute carrier family 23 member 2	Q9EPR4	ENSCINP00000007688.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE22065	SAHH	Adenosylhomocysteinase	P27604	ENSCINP00000011751.3	GO:0004013 GO:0005829 GO:0006730 GO:0016787 GO:0033353
AIPGENE8159	SAHH	Adenosylhomocysteinase	P27604	ENSCINP00000011751.3	GO:0004013 GO:0005829 GO:0006730

					GO:0016787 GO:0033353
AIPGENE6185	SC5A2	Sodium/glucose cotransporter 2	Q923I7	ENSCINP00000014618.3	
AIPGENE15912	SC5A3	Sodium/myo-inositol cotransporter	P53793	ENSCINP00000014742.3	
AIPGENE9202	SENP2	Sentrin-specific protease 2	Q5R7K7	ENSCINP00000020258.3	GO:0004175 GO:0005634 GO:0006508 GO:0008234 GO:0016926 GO:0016929
AIPGENE19729	SLD1	Delta(8)-fatty-acid desaturase	Q43469	ENSCINP00000019998.3	
AIPGENE3185	SNED1	Sushi, nidogen and EGF-like domain-containing protein 1	Q70E20	ENSCINP00000030555.1	GO:0005509 GO:0007160
AIPGENE9034	SRSF4	Serine/arginine-rich splicing factor 4	Q8VE97	ENSCINP00000010229.3	GO:0000166 GO:0003676
AIPGENE9107	SVOP	Synaptic vesicle 2-related protein	Q2XWK0	ENSCINP00000013183.3	GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE12941	SYT9	Synaptotagmin-9	Q86SS6	ENSCINP00000001641.3	GO:0005509 GO:0005544 GO:0005886 GO:0006887 GO:0006906 GO:0016020 GO:0017158 GO:0019905 GO:0030276 GO:0048791 GO:0098793
AIPGENE3467	T53I2	Tumor protein p53-inducible nuclear protein 2	Q8IXH6		
AIPGENE13971	TC1A	Transposable element Tc1 transposase	P03934		
AIPGENE29241	TCNA	Sialidase	P23253	ENSCINP00000024742.2	
AIPGENE18308	THAP9	DNA transposase THAP9	Q9H5L6	ENSCINP00000026647.2	
AIPGENE8999	TNG2	Transport and Golgi organization protein 2 homolog	Q6ICL3		
AIPGENE27609	TNR27	Tumor necrosis factor receptor superfamily member 27	Q8BX35		
AIPGENE5278	TTC28	Tetratricopeptide repeat protein 28	Q80XJ3	ENSCINP00000018424.3	
AIPGENE14672	TYRO	Tyrosinase	P55023	ENSCINP00000008838.3	GO:0008152 GO:0016491 GO:0046872 GO:0055114
AIPGENE3150	USOM5	Uncharacterized skeletal organic matrix protein 5	B8VIU6	ENSCINP00000013341.3	GO:0005344 GO:0006810 GO:0015671 GO:0019825 GO:0020037
AIPGENE16693	VWDE	von Willebrand factor D and EGF domain-containing protein	Q8N2E2	ENSCINP00000026946.2	
AIPGENE25233	WHITE	Protein white	Q05360	ENSCINP00000009488.3	GO:0000166 GO:0005524 GO:0005886 GO:0006727 GO:0016020 GO:0016021 GO:0016887 GO:0042626 GO:0055085
AIPGENE12685	XDH	Xanthine dehydrogenase	P10351	ENSCINP00000031513.1	GO:0003824 GO:0004854 GO:0005506

					GO:0005829 GO:0009055 GO:0009115 GO:0016491 GO:0016614 GO:0016903 GO:0046872 GO:0050660 GO:0051536 GO:0055114
AIPGENE19821	XDH	Xanthine dehydrogenase/oxidase	P80457	ENSCINP00000031513.1	GO:0003824 GO:0004854 GO:0005506 GO:0005829 GO:0009055 GO:0009115 GO:0016491 GO:0016614 GO:0016903 GO:0046872 GO:0050660 GO:0051536 GO:0055114
AIPGENE11300	XLRS1	Retinoschisin	Q9W6R5	ENSCINP00000019787.3	
AIPGENE20620	ZSWM3	Zinc finger SWIM domain-containing protein 3	Q8CFL8	ENSCINP00000016654.3	GO:0003677 GO:0003700 GO:0005634 GO:0006355 GO:0006366 GO:0008270 GO:0071277
AIPGENE10406	Predicted protein	Predicted protein	A0A096M4T5		
AIPGENE10918	Predicted protein	Predicted protein	A7S1V5		
AIPGENE11265	Predicted Protein	Predicted Protein			
AIPGENE11357	Predicted protein	Predicted protein	A7SUN8		
AIPGENE12057	Predicted Protein	Predicted Protein			
AIPGENE12101	Predicted protein	Predicted protein	A7RUA0	ENSCINP00000032263.1	
AIPGENE14041	Predicted Protein	Predicted Protein			
AIPGENE14493	Predicted Protein	Predicted Protein			
AIPGENE16517	Predicted protein	Predicted protein	A7RQD5		
AIPGENE16818	Predicted protein	Predicted protein	A7T0C9		
AIPGENE17124	Predicted Protein	Predicted Protein			
AIPGENE17166	Predicted Protein	Predicted Protein			
AIPGENE17272	Predicted protein	Predicted protein	C3ZUM2		
AIPGENE17491	Predicted protein	Predicted protein	T1J4M1		
AIPGENE18008	Predicted Protein	Predicted Protein			
AIPGENE18011	Predicted Protein	Predicted Protein			
AIPGENE18035	Predicted Protein	Predicted Protein			
AIPGENE19135	Predicted protein	Predicted protein	R7TT16		
AIPGENE19608	Predicted protein	Predicted protein	C3ZHH9		
AIPGENE19944	Predicted Protein	Predicted Protein			
AIPGENE20050	Predicted Protein	Predicted Protein			
AIPGENE20128	Predicted Protein	Predicted Protein			
AIPGENE2077	Predicted protein	Predicted protein	A7RQH4		
AIPGENE21447	Predicted Protein	Predicted Protein			
AIPGENE21467	Predicted Protein	Predicted Protein			
AIPGENE21580	Predicted protein	Predicted protein	C3ZYA3		
AIPGENE23948	Predicted Protein	Predicted Protein			
AIPGENE2407	Predicted protein	Predicted protein	A7SZH2		
AIPGENE25323	Predicted protein	Predicted protein	A7RHU0		
AIPGENE25330	Predicted Protein	Predicted Protein			
AIPGENE26379	Predicted protein	Predicted protein	A7SH51		
AIPGENE26964	Predicted Protein	Predicted Protein			
AIPGENE27013	Predicted protein	Predicted protein	A7SA71		
AIPGENE28315	Predicted protein	Predicted protein	W4XQ99		
AIPGENE28481	Predicted protein	Predicted protein	A7RQ30	ENSCINP00000005729.3	GO:0005737 GO:0006629 GO:0016020 GO:0016021 GO:0016627 GO:0055114
AIPGENE28909	Predicted Protein	Predicted Protein			
AIPGENE29055	Predicted Protein	Predicted Protein			
AIPGENE3576	Predicted protein	Predicted protein	A7S714	ENSCINP00000001358.3	GO:0016020

					GO:0016021
AIPGENE5067	Predicted protein	Predicted protein	A0A015J949	ENSCINP00000034291.1	
AIPGENE6555	Predicted protein	Predicted protein	Q110F0		
AIPGENE6567	Predicted protein	Predicted protein	A7SUI3		
AIPGENE6793	Predicted protein	Predicted protein	A7RZH9		
AIPGENE6924	Predicted Protein	Predicted Protein			
AIPGENE72	Predicted Protein	Predicted Protein			
AIPGENE8511	Predicted protein	Predicted protein	I3KX76		
AIPGENE8559	Predicted Protein	Predicted Protein			
AIPGENE9347	Predicted Protein	Predicted Protein			
AIPGENE9410	Predicted protein	Predicted protein	A0A096M4T5		

Table S1-2. Enriched GO terms for the host DEGs unique to the symbiotic individuals

E. diaphana geneID	Gene name	Discription	UniprotID	*C. intestinalis* geneID	GOID
AIPGENE11786	4CL1	4-coumarate—CoA ligase 1	O24145	ENSCINP00000036514.1	GO:0003824 GO:0008152
AIPGENE527	5HT6R	5-hydroxytryptamine receptor 6	P31388	ENSCINP00000015775.3	
AIPGENE24100	6PGD	6-phosphogluconate dehydrogenase, decarboxylating	P52209	ENSCINP00000031881.1	GO:0004616 GO:0055114
AIPGENE22099	AASS	Alpha-aminoadipic semialdehyde synthase, mitochondrial	A8E657	ENSCINP00000019576.3	GO:0016491 GO:0055114
AIPGENE20833	ABCG4	ATP-binding cassette sub-family G member 4	Q9H172	ENSCINP00000016671.3	GO:0000166 GO:0005524 GO:0005886 GO:0015918 GO:0016020 GO:0016021 GO:0016887 GO:0017127 GO:0030301 GO:0034041
AIPGENE2778	ABHD5	1-acylglycerol-3-phosphate O-acyltransferase ABHD5	Q9DBL9	ENSCINP00000019002.3	GO:0003824 GO:0016787
AIPGENE11605	ACAC	Acetyl-CoA carboxylase	P11029	ENSCINP00000005865.3	GO:0000166 GO:0003824 GO:0003989 GO:0004075 GO:0005524 GO:0006633 GO:0016874 GO:0046872
AIPGENE13427	ACADM	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	Q3SZB4	ENSCINP00000030042.1	GO:0000062 GO:0003995 GO:0005739 GO:0008152 GO:0008470 GO:0009055 GO:0016491 GO:0016627 GO:0033539 GO:0050660 GO:0052890 GO:0055088 GO:0055114
AIPGENE17854	ACEA	Isocitrate lyase	Q9K9H0		
AIPGENE18437	ACES	Acetylcholinesterase	Q92035	ENSCINP00000031848.1	
AIPGENE18552	ACES	Acetylcholinesterase	P22303	ENSCINP00000035547.1	GO:0016787
AIPGENE983	ACES	Acetylcholinesterase	Q92035	ENSCINP00000009596.3	GO:0004104 GO:0005615 GO:0016787 GO:0052689
AIPGENE20290	ACOD	Acyl-CoA desaturase	O62849	ENSCINP00000033408.1	
AIPGENE1750	ACOT4	Acyl-coenzyme A thioesterase 4	Q8BWN8	ENSCINP00000025919.2	GO:0006637 GO:0016790
AIPGENE12816	ACTB	Actin, cytoplasmic 1	P15475	ENSCINP00000001538.3	GO:0000166 GO:0005524
AIPGENE12775	ACTC	Actin, cytoplasmic	P12716	ENSCINP00000001538.3	GO:0000166 GO:0005524
AIPGENE2600	ACTP2	Equinatoxin-2	P61914		
AIPGENE2697	ACTP2	Equinatoxin-2	P61914		
AIPGENE2649	ACTPC	Tenebrosin-C	P61915		
AIPGENE19049	AEBP1	Adipocyte enhancer-binding protein 1	Q640N1	ENSCINP00000019787.3	
AIPGENE25748	AGRD1	Adhesion G-protein coupled receptor D1	A6QLU6	ENSCINP00000011544.3	GO:0004871 GO:0004888 GO:0004930 GO:0005509 GO:0005886 GO:0007155 GO:0007156 GO:0007165 GO:0007166

					GO:0007186 GO:0016020 GO:0016021
AIPGENE20632	AGRIN	Agrin	P25304	ENSCINP00000024218.2	GO:0016020 GO:0016021
AIPGENE19640	AL3B1	Aldehyde dehydrogenase family 3 member B1	P43353	ENSCINP00000001020.3	GO:0004028 GO:0004029 GO:0004030 GO:0006081 GO:0008152 GO:0016491 GO:0016620 GO:0055114
AIPGENE4065	AMDA	Peptidyl-glycine alpha-amidating monooxygenase A	P08478	ENSCINP00000018020.4	GO:0003824 GO:0004497 GO:0005507 GO:0006518 GO:0016020 GO:0016715 GO:0055114
AIPGENE5791	AMID	Amidase	P95896	ENSCINP00000027138.2	GO:0016884
AIPGENE28006	ANAG	Alpha-N-acetylglucosaminidase	P54802	ENSCINP00000024455.2	
AIPGENE23315	AOSL	Allene oxide synthase-lipoxygenase protein	O16025	ENSCINP00000033534.1	GO:0016491 GO:0016702 GO:0046872 GO:0051213 GO:0055114
AIPGENE5921	ARF1	ADP-ribosylation factor 1	P61210	ENSCINP00000035274.1	
AIPGENE18609	ARID2	AT-rich interactive domain-containing protein 2	Q68CP9	ENSCINP00000013045.3	GO:0001664 GO:0005737 GO:0005764 GO:0006629 GO:0006665 GO:0008047 GO:0043085
AIPGENE6303	ARSB	Arylsulfatase B	P15848	ENSCINP00000007178.3	GO:0003824 GO:0008152 GO:0008484
AIPGENE6329	ARSB	Arylsulfatase B	P33727	ENSCINP00000007178.3	GO:0003824 GO:0008152 GO:0008484
AIPGENE12193	ATHL1	Acid trehalase-like protein 1	Q8BP56	ENSCINP00000036547.1	GO:0003824 GO:0005618 GO:0005975 GO:0016787
AIPGENE18100	ATS18	A disintegrin and metalloproteinase with thrombospondin motifs 18	Q8TE60	ENSCINP00000018188.3	GO:0004222 GO:0005578 GO:0006508 GO:0008237 GO:0008270 GO:0031012
AIPGENE19799	ATS6	A disintegrin and metalloproteinase with thrombospondin motifs 6	Q9UKP5	ENSCINP00000014955.3	GO:0004222 GO:0005578 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0031012
AIPGENE7831	ATS6	A disintegrin and metalloproteinase with thrombospondin motifs 6	Q9UKP5	ENSCINP00000014955.3	GO:0004222 GO:0005578 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0031012
AIPGENE3413	ATS7	disintegrin and metalloproteinase with thrombospondin motifs 7	Q1EHB3	ENSCINP00000030146.1	
AIPGENE3533	ATTY	Tyrosine aminotransferase	Q8QZR1	ENSCINP00000000577.3	GO:0003824 GO:0004838 GO:0006520 GO:0008483 GO:0009058

					GO:0009072 GO:0030170
AIPGENE20839	B4GN3	Beta-1,4-N-acetylgalactosaminyltransferase 3	Q6L9W6	ENSCINP00000028502.2	GO:0008376 GO:0032580
AIPGENE14704	BACE	Beta-secretase	W8W138	ENSCINP00000008898.3	GO:0004190 GO:0006508 GO:0008233 GO:0016020 GO:0016021 GO:0016787 GO:0030163
AIPGENE991	BAI3	Brain-specific angiogenesis inhibitor 3	Q80ZF8	ENSCINP00000015858.3	GO:0005509
AIPGENE18881	BAMT	Benzoate carboxyl methyltransferase	Q9FYZ9		
AIPGENE12883	BARH1	BarH-like 1 homeobox protein	Q9BZE3	ENSCINP00000024460.2	GO:0003677 GO:0005634 GO:0006355 GO:0043565
AIPGENE10563	BAT38	BTB and MATH domain-containing protein 38	Q20681	ENSCINP00000033484.1	GO:0004842 GO:0016567 GO:0031463
AIPGENE1323	BCAS3	Breast carcinoma-amplified sequence 3 homolog	Q8CCN5	ENSCINP00000024644.2	
AIPGENE13568	BDH2	3-hydroxybutyrate dehydrogenase type 2	Q3T046	ENSCINP00000003163.3	GO:0003858 GO:0019290 GO:0055072 GO:0055114
AIPGENE9338	BGLR	Beta-glucuronidase	O97524	ENSCINP00000032126.1	GO:0004553 GO:0004566 GO:0005764 GO:0005975 GO:0008152 GO:0016020 GO:0016021 GO:0016787 GO:0016798
AIPGENE9365	BGLR	Beta-glucuronidase	Q4FAT7	ENSCINP00000010614.3	GO:0004553 GO:0004566 GO:0005764 GO:0005975 GO:0008152 GO:0016787 GO:0016798 GO:0043231
AIPGENE27531	BHMT1	Betaine—homocysteine S-methyltransferase 1	Q5M8Z0	ENSCINP00000015488.3	GO:0005737 GO:0005829 GO:0008270 GO:0008898 GO:0009086 GO:0032259 GO:0046872 GO:0047150
AIPGENE1836	BPI	Bactericidal permeability-increasing protein	P17213	ENSCINP00000033806.1	
AIPGENE28055	BTN1	Protein BTN1	Q6CLN9		
AIPGENE8360	C1T9B	Complement C1q and tumor necrosis factor-related protein 9B	B2RNN3	ENSCINP00000032147.1	
AIPGENE2901	CAH2	Carbonic anhydrase 2	Q8UWA5	ENSCINP00000003284.3	GO:0004089 GO:0006730 GO:0008270 GO:0046872
AIPGENE11652	CALUA	Calumenin-A	B5X186	ENSCINP00000025165.2	GO:0005509
AIPGENE18736	CAN8	Calpain-8	Q91VA3	ENSCINP00000015502.3	
AIPGENE24429	CAS4	Short-chain collagen C4 (Fragment)	P18503	ENSCINP00000032915.1	GO:0016020 GO:0016021
AIPGENE24564	CAS4	Short-chain collagen C4 (Fragment)	P18503	ENSCINP00000032915.1	GO:0016020 GO:0016021
AIPGENE22675	CASR	Extracellular calcium-sensing receptor	P35384	ENSCINP00000004243.3	GO:0004871 GO:0004930 GO:0005886 GO:0007165

					GO:0007186 GO:0016020 GO:0016021
AIPGENE16849	CATB	Cathepsin B	A1E295	ENSCINP00000031460.1	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008234 GO:0050790 GO:0051603
AIPGENE26156	CATL	Cathepsin L	Q26636	ENSCINP00000015731.3	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008233 GO:0008234 GO:0016787 GO:0051603
AIPGENE8126	CC125	Coiled-coil domain-containing protein 125	Q5U465		
AIPGENE128	CC134	Coiled-coil domain-containing protein 134	Q9H6E4	ENSCINP00000021785.2	
AIPGENE22347	CD033	UPF0462 protein C4orf33	Q8N1A6	ENSCINP00000032660.1	
AIPGENE13459	CD36	Platelet glycoprotein 4	P70110	ENSCINP00000011713.3	GO:0004872 GO:0005764 GO:0016020 GO:0016021
AIPGENE1298	CD63	CD63 antigen	Q76B49	ENSCINP00000031139.1	
AIPGENE24299	CDAC1	Cytidine and dCMP deaminase domain-containing protein 1	Q9BWV3	ENSCINP00000026466.2	GO:0003824 GO:0004132 GO:0006220 GO:0008270 GO:0016787 GO:0046872
AIPGENE3703	CDCP2	CUB domain-containing protein 2	Q5VXM1	ENSCINP00000032546.1	
AIPGENE1271	CFAD	Counting factor associated protein D	Q54TR1	ENSCINP00000015736.3	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008234 GO:0051603
AIPGENE15465	CGAS	Cyclic GMP-AMP synthase	I3LM39	ENSCINP00000021875.2	
AIPGENE11370	CGL	Cystathionine gamma-lyase	Q55DV9	ENSCINP00000025947.2	GO:0003824 GO:0030170
AIPGENE510	CGL	Cystathionine gamma-lyase	Q58DW2	ENSCINP00000025947.2	GO:0003824 GO:0030170
AIPGENE7524	CGT	2-hydroxyacylsphingosine 1-beta-galactosyltransferase	Q16880		
AIPGENE28896	CHDH	Choline dehydrogenase, mitochondrial	Q8BJ64	ENSCINP00000007524.3	GO:0008812 GO:0016491 GO:0016614 GO:0019285 GO:0050660 GO:0055114
AIPGENE3160	CHST1	Carbohydrate sulfotransferase 1	Q6DBY9	ENSCINP00000022864.2	GO:0001517 GO:0006790 GO:0008146 GO:0016021 GO:0016740
AIPGENE22861	CKAP5	Cytoskeleton-associated protein 5	Q14008	ENSCINP00000015580.3	
AIPGENE22622	CLIC4	Chloride intracellular channel protein 4	Q5R957	ENSCINP00000009622.3	GO:0005254 GO:0005634 GO:0005737 GO:0050780 GO:1902476
AIPGENE7759	CNG	Cyclic nucleotide-gated cation channel	P55934	ENSCINP00000014322.3	GO:0005216 GO:0005249 GO:0005887 GO:0006810 GO:0006811 GO:0006813 GO:0016020

					GO:0016021 GO:0034220 GO:0042391 GO:0055085 GO:0071805
AIPGENE4172	CNN1	Calponin-1	P26932	ENSCINP00000024676.2	GO:0003779 GO:0005516 GO:0031032
AIPGENE20994	CNR2	Cannabinoid receptor 2	P47936	ENSCINP00000008927.3	GO:0004930 GO:0005887 GO:0007166 GO:0007186 GO:0016020 GO:0016021
AIPGENE10422	CO4A1	Collagen alpha-1(IV) chain	P08120	ENSCINP00000016017.3	
AIPGENE10546	CO4A2	Collagen alpha-2(IV) chain	P27393	ENSCINP00000016093.3	GO:0005201 GO:0005578 GO:0005581
AIPGENE25153	CO6A5	Collagen alpha-5(VI) chain	A6H584	ENSCINP00000001942.3	GO:0005509
AIPGENE14423	CO7A1	Collagen alpha-1(VII) chain	Q63870	ENSCINP00000017737.3	GO:0004867 GO:0010951
AIPGENE14130	COLA1	Collagen alpha-1(XXI) chain	Q96P44	ENSCINP00000016093.3	GO:0005201 GO:0005578 GO:0005581
AIPGENE19050	COLA1	Collagen alpha-1(XXI) chain	Q96P44	ENSCINP00000016017.3	
AIPGENE4529	COLL2	Collagen-like protein 2	Q5UQ13	ENSCINP00000017221.3	
AIPGENE20999	COLL6	Collagen-like protein 6	Q5UQ50	ENSCINP00000001599.3	
AIPGENE15090	COLQ	Acetylcholinesterase collagenic tail peptide	Q03637	ENSCINP00000014311.3	GO:0005201 GO:0005581
AIPGENE1629	COTL1	Coactosin-like protein	Q2HJ57	ENSCINP00000017953.3	
AIPGENE23049	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	O73853	ENSCINP00000019654.3	GO:0004497 GO:0005506 GO:0008395 GO:0016020 GO:0016021 GO:0016491 GO:0016705 GO:0020037 GO:0046872 GO:0055114
AIPGENE23070	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	O73853	ENSCINP00000022528.2	
AIPGENE23109	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	O73853	ENSCINP00000018706.3	
AIPGENE27064	CP1A1	Cytochrome P450 1A1	O42457	ENSCINP00000012842.3	GO:0004497 GO:0005506 GO:0016020 GO:0016021 GO:0016491 GO:0016705 GO:0020037 GO:0043231 GO:0046872 GO:0055114
AIPGENE5640	CP24A	1,25-dihydroxyvitamin D(3) 24-hydroxylase, mitochondrial	Q64441	ENSCINP00000033359.1	
AIPGENE27007	CP3AO	Cytochrome P450 3A24	Q29496	ENSCINP00000036138.1	
AIPGENE28157	CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial	P31327	ENSCINP00000018277.3	GO:0000050 GO:0003824 GO:0004070 GO:0004087 GO:0004088 GO:0004151 GO:0005524 GO:0005737 GO:0006207 GO:0006526 GO:0006807 GO:0046872
AIPGENE21331	CPT1A	Carnitine O-palmitoyltransferase 1, liver isoform	P50416	ENSCINP00000007072.3	GO:0016020 GO:0016021 GO:0016746
AIPGENE13554	CREA	Creatinase	P38487		
AIPGENE26823	CTHR1	Collagen triple helix repeat-	Q96CG8	ENSCINP00000015636.3	

		containing protein 1			
AIPGENE13178	CTP5A	Contactin-associated protein like 5-1	Q0V8T9	ENSCINP00000010540.3	
AIPGENE24260	CTRB	Chymotrypsin B	P80646	ENSCINP00000029043.2	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE5888	CTRB	Chymotrypsinogen B	P00767	ENSCINP00000031261.1	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE6698	CTRB2	Chymotrypsinogen B2	Q6GPI1	ENSCINP00000029043.2	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE10951	CTX1	Toxin CrTX-A	Q9GV72		
AIPGENE17825	CTX1	Toxin CrTX-A	Q9GV72		
AIPGENE17834	CTX1	Toxin CrTX-A	Q9GV72		
AIPGENE17675	CTXA	Toxin CaTX-A	Q9GNN8		
AIPGENE13588	CUBN	Cubilin	O70244	ENSCINP00000007799.3	GO:0004222 GO:0005509 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787 GO:0046872
AIPGENE5556	CYS2	Probable serine-O-acetyltransferase cys2	Q10341	ENSCINP00000033915.1	GO:0009001 GO:0016787 GO:0019344
AIPGENE22204	CYSP	Cathepsin B-like cysteine proteinase	P25792	ENSCINP00000019058.3	GO:0004177 GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0006919 GO:0006955 GO:0008233 GO:0008234 GO:0008656 GO:0010952 GO:0016787 GO:0051603 GO:2001235
AIPGENE26801	DEFM	Peptide deformylase, mitochondrial	Q9HBH1		
AIPGENE6491	DHCR7	7-dehydrocholesterol reductase	Q7SXF1	ENSCINP00000014788.3	GO:0006695 GO:0016020 GO:0016021 GO:0016628 GO:0030176 GO:0047598 GO:0055114
AIPGENE24026	DHE3	Glutamate dehydrogenase	P94598	ENSCINP00000019611.3	GO:0006520 GO:0016491 GO:0016639 GO:0055114
AIPGENE27618	DHE3	Glutamate dehydrogenase	P94598	ENSCINP00000019611.3	GO:0006520 GO:0016491 GO:0016639 GO:0055114
AIPGENE25704	DIAC	Di-N-acetylchitobiase	Q01459	ENSCINP00000035855.1	
AIPGENE20433	DIRA2	GTP-binding protein Di-Ras2	Q5R6S2	ENSCINP00000015095.3	GO:0000166 GO:0005525 GO:0005622 GO:0007165 GO:0007264 GO:0016020
AIPGENE20966	DIRC2	Disrupted in renal carcinoma protein 2 homolog	Q6GNV7	ENSCINP00000023309.2	
AIPGENE15640	DMBT1	Deleted in malignant brain	Q9UGM3	ENSCINP00000015917.3	GO:0004252

		tumors 1 protein			GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE5579	DNAJ1	DnaJ protein homolog 1	Q24133	ENSCINP00000012873.3	GO:0006457 GO:0051082
AIPGENE26101	DNJB4	DnaJ homolog subfamily B member 4	Q2KIT4	ENSCINP00000002844.3	GO:0006457 GO:0051082
AIPGENE3208	DPF1	Zinc finger protein neuro-d4	Q9QX66	ENSCINP00000033757.1	GO:0008270 GO:0046872
AIPGENE17142	DPP2	Dipeptidyl peptidase 2	Q9EPB1	ENSCINP00000014279.3	GO:0004185 GO:0006508 GO:0008236 GO:0008239 GO:0031982
AIPGENE24798	DRD2B	D(2) dopamine receptor B (Fragment)	P34973		
AIPGENE3939	DSCAM	Down syndrome cell adhesion molecule	O60469	ENSCINP00000015860.3	GO:0005509
AIPGENE29265	DTX3L	E3 ubiquitin-protein ligase DTX3L	Q8TDB6	ENSCINP00000006511.3	GO:0008270 GO:0046872
AIPGENE24155	DYH3	Dynein heavy chain 3, axonemal	Q8BW94	ENSCINP00000031667.1	GO:0003777 GO:0007018
AIPGENE12845	ECE1	Endothelin-converting enzyme 1	P42893	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE8891	EDIL3	EGF-like repeat and discoidin I-like domain-containing protein 3	O35474	ENSCINP00000006532.3	GO:0005509
AIPGENE25107	EIF2A	Eukaryotic translation initiation factor 2A	Q4QRJ7	ENSCINP00000005103.3	GO:0003743 GO:0006412 GO:0006413
AIPGENE8234	EIF3D	Eukaryotic translation initiation factor 3 subunit D	A7SMR1	ENSCINP00000008206.1	
AIPGENE10025	ELOV4	Elongation of very long chain fatty acids protein 4	Q9GZR5	ENSCINP00000018836.3	GO:0006629 GO:0006631 GO:0006633 GO:0016020 GO:0016021 GO:0016740
AIPGENE24295	EMX1	Homeobox protein EMX1	Q04741	ENSCINP00000036336.1	GO:0003677 GO:0005634 GO:0006355 GO:0043565
AIPGENE14033	ENOX1	Ecto-NOX disulfide-thiol exchanger 1	Q8TC92		
AIPGENE25794	ENOX1	Ecto-NOX disulfide-thiol exchanger 1	Q8TC92		
AIPGENE3034	EPDR1	Mammalian ependymin-related protein 1	Q9N0C7	ENSCINP00000015708.3	GO:0005509 GO:0005576 GO:0007160
AIPGENE14889	EPHB2	Ephrin type-B receptor 2	P54763	ENSCINP00000002487.3	GO:0000166 GO:0004672 GO:0004713 GO:0005003 GO:0005524 GO:0005887 GO:0006468 GO:0007169 GO:0016020 GO:0016021 GO:0016301 GO:0016310 GO:0016740 GO:0018108 GO:0048013
AIPGENE8576	EPHB2	Ephrin type-B receptor 2	P54763	ENSCINP00000002487.3	GO:0000166 GO:0004672 GO:0004713 GO:0005003 GO:0005524 GO:0005887 GO:0006468 GO:0007169

					GO:0016020 GO:0016021 GO:0016301 GO:0016310 GO:0016740 GO:0018108 GO:0048013
AIPGENE20366	EQST	Equistatin	P81439	ENSCINP00000020167.3	
AIPGENE26107	ERI2	ERI1 exoribonuclease 2	A8K979	ENSCINP00000012755.3	GO:0000467 GO:0003676
AIPGENE18149	ERMP1	Endoplasmic reticulum metallopeptidase 1	Q3UVK0	ENSCINP00000035901.1	GO:0016020 GO:0016021
AIPGENE391	ERP44	Endoplasmic reticulum resident protein 44	Q9D1Q6	ENSCINP00000002784.3	GO:0003756 GO:0005783 GO:0006457 GO:0034976 GO:0045454
AIPGENE20815	FA26F	Protein FAM26F	Q8C9E8	ENSCINP00000007979.3	GO:0005261 GO:0005887 GO:0016020 GO:0016021 GO:0034220 GO:0098655
AIPGENE19098	FA46C	Protein FAM46C	Q7ZUP1	ENSCINP00000026866.2	
AIPGENE26056	FAAH	Fatty acid amide hydrolase	Q7XJJ7	ENSCINP00000014515.3	
AIPGENE11624	FABG	3-oxoacyl-[acyl-carrier-protein] reductase FabG	Q9X248	ENSCINP00000020171.3	
AIPGENE14819	FAD5	Delta(5) fatty acid desaturase	O74212	ENSCINP00000004431.3	GO:0016020 GO:0016021 GO:0020037 GO:0046872
AIPGENE17789	FAMT	Probable fatty acid methyltransferase Rv3720	O69687		
AIPGENE8227	FAS	Fatty acid synthase	P12785	ENSCINP00000036430.1	GO:0003824 GO:0008152 GO:0009058 GO:0016491 GO:0016740 GO:0016788 GO:0055114
AIPGENE11718	FAXC	Failed axon connections homolog	F7E235	ENSCINP00000025230.2	GO:0004364 GO:0005737 GO:0006749
AIPGENE4649	FB252	F-box protein At5g06550	Q67XX3	ENSCINP00000015483.3	
AIPGENE18349	FBN1	Fibrillin-1	P35555	ENSCINP00000010385.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE8199	FBN1	Fibrillin-1	P98133	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE11705	FBN2	Fibrillin-2	P35556	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE11707	FBN2	Fibrillin-2	P35556	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE13620	FBN2	Fibrillin-2	P35556	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE27824	FBP1	Fibropellin-1	P10079	ENSCINP00000014588.3	
AIPGENE26088	FGF1	Fibroblast growth factor 1 (Fragment)	Q7SIF8	ENSCINP00000032792.1	GO:0008083
AIPGENE1796	FGFR	Fibroblast growth factor receptor	Q4H3K6	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE4615	FGFR	Fibroblast growth factor receptor	Q86PM4	ENSCINP00000026874.2	

AIPGENE16047	FGFR1	Fibroblast growth factor receptor 1	P22182	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE22541	FGFR2	Fibroblast growth factor receptor 2	Q91286	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE7347	FGFR4	Fibroblast growth factor receptor 4	Q90330	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE3281	FKBP4	46 kDa FK506-binding nuclear protein	Q26486		
AIPGENE12192	FMO2	Dimethylaniline monooxygenase [N-oxide-forming] 2	Q8HZ70	ENSCINP00000011417.3	GO:0004497 GO:0004499 GO:0005783 GO:0005789 GO:0016020 GO:0016021 GO:0016491 GO:0017144 GO:0031090 GO:0043231 GO:0050660 GO:0050661 GO:0055114
AIPGENE21768	FMO5	Dimethylaniline monooxygenase [N-oxide-forming] 5	Q8K4C0	ENSCINP00000011417.3	GO:0004497 GO:0004499 GO:0005783 GO:0005789 GO:0016020 GO:0016021 GO:0016491 GO:0017144 GO:0031090 GO:0043231 GO:0050660 GO:0050661 GO:0055114
AIPGENE5457	FOLH1	Glutamate carboxypeptidase 2	P70627	ENSCINP00000011038.3	GO:0016020 GO:0016021
AIPGENE13402	FOSX	Transforming protein v-Fos/v-Fox	P29176	ENSCINP00000034304.1	
AIPGENE3308	FOSX	Transforming protein v-Fos/v-Fox	P29176	ENSCINP00000005786.3	GO:0000978 GO:0000982 GO:0003677 GO:0003700 GO:0005634 GO:0006355 GO:0006357 GO:0043565
AIPGENE26282	FOXO	Forkhead box protein O	O16850	ENSCINP00000035316.1	GO:0003677 GO:0003700 GO:0005634 GO:0006351

					GO:0006355 GO:0043565
AIPGENE20026	FRIS	Soma ferritin	P42577	ENSCINP00000035294.1	
AIPGENE14754	FRRS1	Putative ferric-chelate reductase 1	Q6INU7	ENSCINP00000017029.3	
AIPGENE563	FRRS1	Ferric-chelate reductase 1	Q8K385	ENSCINP00000017029.3	
AIPGENE9887	FSTL5	Follistatin-related protein 5	Q8N475	ENSCINP00000025719.2	
AIPGENE14536	FUCO_BRAFL	Alpha-L-fucosidase2	C3YWU0	ENSCINP00000006232.3	GO:0004560 GO:0005975 GO:0006004
AIPGENE16012	FUCO_BRAFL	Alpha-L-fucosidase2	C3YWU0	ENSCINP00000015524.3	
AIPGENE24268	FUCO_BRAFL	Alpha-L-fucosidase2	C3YWU0	ENSCINP00000015524.3	
AIPGENE15521	FXDC2	Fatty acid hydroxylase domain-containing protein 2	Q96IV6	ENSCINP00000010642.3	GO:0005506 GO:0006633 GO:0016020 GO:0016021 GO:0016491 GO:0055114
AIPGENE12778	GALK1	Galactokinase	P51570	ENSCINP00000010220.3	GO:0000166 GO:0004335 GO:0005524 GO:0005737 GO:0006012 GO:0008152 GO:0016301 GO:0016310 GO:0016740 GO:0016773 GO:0046835
AIPGENE10000	GALK2	N-acetylgalactosamine kinase	Q5XIG6	ENSCINP00000028204.2	GO:0000166 GO:0004335 GO:0005524 GO:0005737 GO:0006012 GO:0008152 GO:0016301 GO:0016310 GO:0016740 GO:0016773 GO:0046835
AIPGENE15776	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000001814.3	GO:0005576
AIPGENE17511	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000001814.3	GO:0005576
AIPGENE1512	GBA3	Cytosolic beta-glucosidase	Q5RF65	ENSCINP00000010342.3	GO:0004553 GO:0005975 GO:0008422 GO:1901657
AIPGENE1160	GBP3	Guanylate-binding protein 3	Q9H0R5	ENSCINP00000032288.1	
AIPGENE23318	GBP5	Guanylate-binding protein 5	Q8CFB4		
AIPGENE5337	GCNT4	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 4	Q71SG7	ENSCINP00000024350.2	
AIPGENE6070	GDF11	Growth/differentiation factor 11	Q9Z1W4	ENSCINP00000010102.3	GO:0005125 GO:0005160 GO:0005576 GO:0005615 GO:0008083 GO:0010862 GO:0030509 GO:0040007 GO:0042981 GO:0043408 GO:0048468 GO:0060395
AIPGENE23694	GGLO	L-gulonolactone oxidase	P58710		
AIPGENE9468	GGT	Gamma-glutamyltranspeptidase	P54422	ENSCINP00000001002.3	GO:0003840 GO:0006749
AIPGENE8953	GLHR	Probable glycoprotein hormone G-protein coupled receptor	P35409	ENSCINP00000030027.1	GO:0004930 GO:0005887 GO:0007186 GO:0007189

					GO:0007190 GO:0008528 GO:0009755 GO:0016020 GO:0016021 GO:0016500
AIPGENE5590	GLI1	Zinc finger protein GLI1 (Fragment)	P55878	ENSCINP00000022660.2	GO:0003676 GO:0006355 GO:0007224 GO:0035301 GO:0046872
AIPGENE3097	GNPAT	Dihydroxyacetone phosphate acyltransferase	O15228	ENSCINP00000018520.3	GO:0008152 GO:0008374 GO:0016746 GO:0044255
AIPGENE182	GNS	N-acetylglucosamine-6-sulfatase	P50426	ENSCINP00000019163.3	
AIPGENE5134	GPSM2	G-protein-signaling modulator 2	P81274	ENSCINP00000018424.3	
AIPGENE26885	GPX1	Glutathione peroxidase 1	P11352	ENSCINP00000002066.3	GO:0004601 GO:0004602 GO:0006979 GO:0016491 GO:0055114 GO:0098869
AIPGENE1907	GRM8	Metabotropic glutamate receptor 8	P47743	ENSCINP00000025414.2	
AIPGENE14421	GRN	Granulins	P23785	ENSCINP00000023895.2	
AIPGENE26283	GRP10	Glycine-rich RNA-binding protein 10	Q05966	ENSCINP00000031642.1	
AIPGENE20480	GST	Glutathione S-transferase	P46428	ENSCINP00000000153.3	
AIPGENE2706	GTR8	Solute carrier family 2, facilitated glucose transporter member 8	Q9NY64	ENSCINP00000002435.3	GO:0005215 GO:0005351 GO:0005355 GO:0005887 GO:0006810 GO:0015992 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0035428 GO:0046323 GO:0055085 GO:1904659
AIPGENE6170	GTR8	Solute carrier family 2, facilitated glucose transporter member 8	Q9JJZ1	ENSCINP00000019539.3	GO:0005351 GO:0005355 GO:0005887 GO:0006810 GO:0015992 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0035428 GO:0046323 GO:0055085 GO:1904659
AIPGENE1127	H17B6	17-beta-hydroxysteroid dehydrogenase type 6	Q3T001	ENSCINP00000015221.3	
AIPGENE15940	HBP1	HMG box-containing protein 1	Q2KJ34	ENSCINP00000015677.3	GO:0000122 GO:0005634
AIPGENE25956	HEBP2	Heme-binding protein 2	Q9WU63	ENSCINP00000019979.3	
AIPGENE25995	HEBP2	Heme-binding protein 2	Q9WU63	ENSCINP00000003777.3	
AIPGENE16319	HEMH	Ferrochelatase, mitochondrial	O57478		
AIPGENE20289	HEX	Beta-hexosaminidase	Q04786	ENSCINP00000008969.3	GO:0004553 GO:0004563 GO:0005975
AIPGENE546	HMCN1	Hemicentin-1	Q96RW7	ENSCINP00000015858.3	GO:0005509
AIPGENE16615	HMCN2	Hemicentin-2	A2AJ76	ENSCINP00000015860.3	GO:0005509
AIPGENE8469	HMCN2	Hemicentin-2	A2AJ76	ENSCINP00000001942.3	GO:0005509
AIPGENE28493	HNF4G	Hepatocyte nuclear factor 4-gamma	Q14541	ENSCINP00000019457.3	
AIPGENE14617	HOIL1	RanBP-type and C3HC4-type	E6ZIJ1	ENSCINP00000033831.1	GO:0008270

		zinc finger-containing protein 1			GO:0046872
AIPGENE24476	HPPD	4-hydroxyphenylpyruvate dioxygenase	Q5BKL0	ENSCINP00000020387.3	GO:0003868 GO:0009072 GO:0016701 GO:0046872 GO:0055114
AIPGENE24739	HRH2	Histamine H2 receptor	P47747	ENSCINP00000009895.3	GO:0004930 GO:0007186 GO:0016020 GO:0016021
AIPGENE18790	HS71A	Heat shock 70 kDa protein 1A	Q61696	ENSCINP00000011757.3	GO:0000166 GO:0005524
AIPGENE21738	HSP7C	Heat shock cognate 71 kDa protein	Q71U34	ENSCINP00000019652.3	
AIPGENE5814	HSP7C	Heat shock cognate 71 kDa protein	Q71U34	ENSCINP00000019652.3	
AIPGENE27608	HSP97	97 kDa heat shock protein	Q94738	ENSCINP00000001654.3	GO:0000166 GO:0005524
AIPGENE28007	HTOMT	Tabersonine 16-O-methyltransferase	B0EXJ8		
AIPGENE17792	HUTH	Histidine ammonia-lyase	P42357	ENSCINP00000013857.3	GO:0003824 GO:0004397 GO:0005737 GO:0006547 GO:0006548 GO:0016829 GO:0016841 GO:0019556 GO:0019557
AIPGENE26189	HYES	Bifunctional epoxide hydrolase 2	Q6Q2C2	ENSCINP00000028721.2	GO:0003824 GO:0008152 GO:0016787
AIPGENE18873	I23O2	Indoleamine 2,3-dioxygenase 2	Q8R0V5		
AIPGENE1254	IF	Gastric intrinsic factor	P17267		
AIPGENE11903	IF2	Translation initiation factor IF-2	B3EAE7		
AIPGENE13893	IF4A1	Eukaryotic initiation factor 4A-I	P60843	ENSCINP00000005836.3	GO:0000166 GO:0003676 GO:0004004 GO:0004386 GO:0005524 GO:0006413 GO:0010468 GO:0010501 GO:0016787
AIPGENE3468	IFI44	Interferon-induced protein 44	Q8TCB0	ENSCINP00000019144.3	
AIPGENE8988	IFI44	Interferon-induced protein 44	Q8TCB0		
AIPGENE10180	IFIH1	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	ENSCINP00000031621.1	
AIPGENE18569	IFIH1	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	ENSCINP00000031621.1	
AIPGENE14622	IIGP5	Interferon-inducible GTPase 5	Q6NXR0		
AIPGENE14629	IIGP5	Interferon-inducible GTPase 5	Q6NXR0		
AIPGENE25676	INO1A	Inositol-3-phosphate synthase 1-A	Q7ZXY0		
AIPGENE20336	INSI2	Insulin-induced gene 2 protein	Q91WG1		
AIPGENE15014	IOD1	Type I iodothyronine deiodinase	O42449	ENSCINP00000035404.1	
AIPGENE19232	IRF2	Interferon regulatory factor 2	Q98925	ENSCINP00000005767.3	GO:0000975 GO:0003700 GO:0006355
AIPGENE24707	IRF2	Interferon regulatory factor 2	P14316	ENSCINP00000005767.3	GO:0000975 GO:0003700 GO:0006355
AIPGENE1568	IRF8	Interferon regulatory factor 8	Q90871	ENSCINP00000001474.3	GO:0000975
AIPGENE20436	JUN	Transcription factor AP-1	P54864	ENSCINP00000018871.3	
AIPGENE1601	K1468	LisH domain and HEAT repeat-containing protein KIAA1468 homolog	Q6P6Y1	ENSCINP00000015377.3	
AIPGENE27863	KARG	Arginine kinase	O15990	ENSCINP00000017993.3	
AIPGENE9229	KDM8	Lysine-specific demethylase 8	Q497B8	ENSCINP00000008667.3	
AIPGENE7287	KLC	Kinesin light chain	P46824	ENSCINP00000034474.1	
AIPGENE9819	KLC	Kinesin light chain	P46822	ENSCINP00000034474.1	

AIPGENE28799	KLF10	Krueppel-like factor 10	O08876	ENSCINP00000012549.3	GO:0003676 GO:0046872
AIPGENE26876	KLHL3	Kelch-like protein 3	E0CZ16	ENSCINP00000008089.3	GO:0004842 GO:0031463 GO:0042787
AIPGENE27005	KLHL4	Kelch-like protein 4	Q9C0H6	ENSCINP00000023174.2	GO:0004842 GO:0016567 GO:0031463
AIPGENE13828	LACS8	Long chain acyl-CoA synthetase 8	Q9SJD4	ENSCINP00000005887.3	GO:0003824 GO:0008152
AIPGENE17012	LAMA4	Laminin subunit alpha-4	Q16363	ENSCINP00000016161.3	
AIPGENE15997	LAMB2	Laminin subunit beta-2	Q61292	ENSCINP00000013062.3	
AIPGENE27242	LAMP1	Lysosome-associated membrane glycoprotein 1	Q05204	ENSCINP00000026601.2	
AIPGENE22791	LAT3	Large neutral amino acids transporter small subunit 3	Q8BSM7	ENSCINP00000036491.1	GO:0016020 GO:0016021
AIPGENE23106	LGMN	Legumain	Q95M12	ENSCINP00000006801.3	GO:0004197 GO:0005773 GO:0006508 GO:0006624 GO:0008233 GO:0051603
AIPGENE27537	LIN41	E3 ubiquitin-protein ligase TRIM71	Q1PRL4	ENSCINP00000023310.2	
AIPGENE26593	LIP1	Lipase ZK262.3	Q9XTR8		
AIPGENE564	LIPR1	Inactive pancreatic lipase-related protein 1	Q5BKQ4	ENSCINP00000014633.3	GO:0004806 GO:0005576 GO:0006629 GO:0052689
AIPGENE8872	LIPR2	Pancreatic lipase-related protein 2	P54318	ENSCINP00000014633.3	GO:0004806 GO:0005576 GO:0006629 GO:0052689
AIPGENE22220	LOX5	Arachidonate 5-lipoxygenase	P48999	ENSCINP00000015881.3	GO:0005506 GO:0016491 GO:0016702 GO:0046872 GO:0051213 GO:0055114
AIPGENE3931	LRIG3	Leucine-rich repeats and immunoglobulin-like domains protein 3	Q6P1C6	ENSCINP00000030254.1	
AIPGENE22337	LRRK2	Leucine-rich repeat serine/threonine-protein kinase 2	Q5S006	ENSCINP00000025708.2	GO:0005525 GO:0005622 GO:0007264 GO:0016787
AIPGENE13786	LSD2	Lipid storage droplets surface-binding protein 2	Q9VXY7		
AIPGENE13692	LYAG	Lysosomal alpha-glucosidase	P10253	ENSCINP00000008112.3	GO:0003824 GO:0004553 GO:0005975 GO:0008152 GO:0016787 GO:0016798 GO:0030246
AIPGENE20624	LYOX	Protein-lysine 6-oxidase	P45845	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE6132	MA2B1	Lysosomal alpha-mannosidase	Q60HE9	ENSCINP00000003992.3	GO:0003824 GO:0004553 GO:0004559 GO:0005975 GO:0006013 GO:0006517 GO:0008152 GO:0008270 GO:0015923 GO:0016787 GO:0016798 GO:0030246 GO:0046872

AIPGENE6148	MA2B1	Lysosomal alpha-mannosidase	Q60HE9	ENSCINP00000003992.3	GO:0003824 GO:0004553 GO:0004559 GO:0005975 GO:0006013 GO:0006517 GO:0008152 GO:0008270 GO:0015923 GO:0016787 GO:0016798 GO:0030246 GO:0046872
AIPGENE26161	MAF	Transcription factor Maf	Q98UK4	ENSCINP00000002531.3	GO:0001228 GO:0003677 GO:0003700 GO:0005634 GO:0006351 GO:0006355 GO:0006357 GO:0006366 GO:0043565 GO:0045944
AIPGENE18937	MAFG	Transcription factor MafG	Q76MX4	ENSCINP00000002531.3	GO:0001228 GO:0003677 GO:0003700 GO:0005634 GO:0006351 GO:0006355 GO:0006357 GO:0006366 GO:0043565 GO:0045944
AIPGENE133	MATN2	Matrilin-2	O08746	ENSCINP00000001942.3	GO:0005509
AIPGENE2489	MATN2	Matrilin-2	O00339	ENSCINP00000014717.3	
AIPGENE11650	MB212	Protein mab-21-like 2	Q8UUZ1	ENSCINP00000021875.2	
AIPGENE16732	MB212	Protein mab-21-like 2	Q8UUZ1	ENSCINP00000021875.2	
AIPGENE4828	MCA7	Metacaspase-7	Q6XPT5		
AIPGENE15953	MCAT	Mitochondrial carnitine/acylcarnitine carrier protein	Q9Z2Z6	ENSCINP00000016773.3	GO:0003735 GO:0006412 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE15960	MCAT	Mitochondrial carnitine/acylcarnitine carrier protein	Q9Z2Z6	ENSCINP00000016773.3	GO:0003735 GO:0006412 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE9553	MET7B	Methyltransferase-like protein 7B	Q6UX53	ENSCINP00000027827.2	GO:0005737 GO:0005829 GO:0008757 GO:0030791 GO:0032259
AIPGENE2103	MFGM	Lactadherin	P70490	ENSCINP00000006532.3	GO:0005509
AIPGENE20554	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE20557	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE26741	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE649	MFSD1	Major facilitator superfamily domain-containing protein 1	Q32LQ6	ENSCINP00000000792.3	GO:0016020 GO:0016021 GO:0055085
AIPGENE14711	MFSD6	Major facilitator superfamily domain-containing protein 6	Q8CBH5	ENSCINP00000006447.3	GO:0016020 GO:0016021
AIPGENE4727	MFSD8	Major facilitator superfamily domain-containing protein 8	Q6GPQ3	ENSCINP00000007756.3	GO:0016020 GO:0016021 GO:0055085
AIPGENE5656	MGLL	Monoglyceride lipase	Q99685	ENSCINP00000020690.3	
AIPGENE14950	MINP1	Multiple inositol polyphosphate phosphatase 1	Q9Z2L6	ENSCINP00000036092.1	GO:0003993 GO:0016311 GO:0016791

AIPGENE815	MIOX	Inositol oxygenase	Q4V8T0	ENSCINP00000009005.3	GO:0005506 GO:0005737 GO:0019310 GO:0050113 GO:0055114
AIPGENE15748	MIP	Lens fiber major intrinsic protein	Q6J8I9	ENSCINP00000011466.3	GO:0005215 GO:0005887 GO:0006810 GO:0006833 GO:0009992 GO:0015250 GO:0015254 GO:0015793 GO:0016020 GO:0016021 GO:0034220
AIPGENE28801	MLRP1	MAM and LDL-receptor class A domain-containing protein 1 (Fragment)	B3EWZ5	ENSCINP00000026839.2	
AIPGENE28176	MLRP2	MAM and LDL-receptor class A domain-containing protein 2 (Fragment)	B3EWZ6	ENSCINP00000036530.1	GO:0016020
AIPGENE19158	MOG2A	2-acylglycerol O-acyltransferase 2-A	Q2KHS5	ENSCINP00000004063.3	GO:0016747
AIPGENE20952	MOG2A	2-acylglycerol O-acyltransferase 2-A	Q2KHS5	ENSCINP00000004063.3	GO:0016747
AIPGENE1928	MOT10	Monocarboxylate transporter 10	Q3U9N9	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE2143	MOT10	Monocarboxylate transporter 10	A1L1W9	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE1268	MOT5	Monocarboxylate transporter 5	Q8R0M8	ENSCINP00000026236.2	
AIPGENE3051	MOT8	Monocarboxylate transporter 8	Q8K1P8	ENSCINP00000026773.2	GO:0005887 GO:0015129 GO:0016020 GO:0016021 GO:0035879 GO:0055085
AIPGENE3158	MOT8	Monocarboxylate transporter 8	Q8K1P8	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE21866	MOXD1	DBH-like monooxygenase protein 1 homolog	Q5TZ24	ENSCINP00000006004.3	GO:0003824 GO:0004497 GO:0005507 GO:0016715 GO:0055114
AIPGENE15605	MPEG1	Macrophage-expressed gene 1 protein	Q9WV57		
AIPGENE1799	MRC1	Macrophage mannose receptor 1	Q61830	ENSCINP00000004176.3	GO:0016020 GO:0016021
AIPGENE14371	MRLCA	Myosin regulatory light chain RLC-A	P13832	ENSCINP00000035076.1	
AIPGENE22589	MRP4	Multidrug resistance-associated protein 4	O15439	ENSCINP00000015473.3	GO:0000166 GO:0005524 GO:0006810 GO:0016020 GO:0016021 GO:0016887 GO:0042626 GO:0055085
AIPGENE25331	MRP4	Multidrug resistance-associated protein 4	O15439	ENSCINP00000015473.3	GO:0000166 GO:0005524 GO:0006810 GO:0016020 GO:0016021 GO:0016887

					GO:0042626 GO:0055085
AIPGENE28235	MRP5	Multidrug resistance-associated protein 5	Q9QYM0	ENSCINP00000013413.3	GO:0000166 GO:0005215 GO:0005524 GO:0006810 GO:0016020 GO:0016021 GO:0016887 GO:0042626 GO:0055085
AIPGENE5829	MTPN	Myotrophin	Q91955	ENSCINP00000012849.2	
AIPGENE25437	MUC1	Integumentary mucin C.1 (Fragment)	Q05049	ENSCINP00000035374.1	
AIPGENE4554	MUC2	Mucin-2 (Fragment)	Q62635	ENSCINP00000035305.1	
AIPGENE14373	MYO3A	Myosin-IIIa	Q8NEV4	ENSCINP00000010911.3	GO:0000166 GO:0004672 GO:0004702 GO:0005524 GO:0005737 GO:0006468 GO:0023014
AIPGENE8101	MYOF	Myoferlin	B3DLH6	ENSCINP00000013715.3	GO:0016020 GO:0016021
AIPGENE25325	MYPH	Myophilin	Q24799	ENSCINP00000025754.2	
AIPGENE4205	MYPH	Myophilin	Q24799	ENSCINP00000025754.2	
AIPGENE3953	N2F1A	Nuclear receptor subfamily 2 group F member 1-A	Q06725	ENSCINP00000004102.3	GO:0003677 GO:0003700 GO:0003707 GO:0004879 GO:0005634 GO:0006351 GO:0006355 GO:0008270 GO:0030522 GO:0043401 GO:0043565 GO:0046872
AIPGENE13027	NAGAB	Alpha-N-acetylgalactosaminidase	Q90744	ENSCINP00000031972.1	GO:0003824 GO:0004553 GO:0004557 GO:0005737 GO:0005975 GO:0008152 GO:0016139 GO:0016787 GO:0016798
AIPGENE13031	NAGAB	Alpha-N-acetylgalactosaminidase	Q90744	ENSCINP00000031972.1	GO:0003824 GO:0004553 GO:0004557 GO:0005737 GO:0005975 GO:0008152 GO:0016139 GO:0016787 GO:0016798
AIPGENE24074	NARPB	Notch-regulated ankyrin repeat-containing protein B	Q7T3X9	ENSCINP00000008656.3	GO:0007165 GO:0019208 GO:0019901 GO:0050790
AIPGENE5975	NAS15	Zinc metalloproteinase nas-15	P55115	ENSCINP00000017083.3	GO:0004222 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016020 GO:0016787 GO:0046872
AIPGENE17549	NAS4	Zinc metalloproteinase nas-4	P55112	ENSCINP00000030500.1	GO:0004222 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787

					GO:0046872
AIPGENE8386	NAT16	Probable N-acetyltransferase 16	Q8N8M0		
AIPGENE25378	NCA11	Neural cell adhesion molecule 1-A	P16170	ENSCINP00000015860.3	GO:0005509
AIPGENE24468	NCL1	B-box type zinc finger protein ncl-1	P34611		
AIPGENE5315	NEP	Neprilysin	P08049	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE15229	NETR	Neurotrypsin	Q5G265	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE25764	NETR	Neurotrypsin	O08762	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE18582	NFH	Neurofilament heavy polypeptide	P19246		
AIPGENE1175	NGB	Neuroglobin	Q99JA8	ENSCINP00000015555.3	GO:0005344 GO:0006810 GO:0015671 GO:0019825 GO:0020037
AIPGENE5536	NKX62	Homeobox protein Nkx-6.2	Q9C056	ENSCINP00000023047.2	GO:0003677 GO:0005634 GO:0006355 GO:0043565
AIPGENE22109	NNMT	Nicotinamide N-methyltransferase	P40261		
AIPGENE5531	NPC1	Niemann-Pick C1 protein	P56941	ENSCINP00000008700.3	GO:0005319 GO:0006869 GO:0016020 GO:0016021
AIPGENE20032	NPC2	Epididymal secretory protein E1	P79345	ENSCINP00000030926.1	
AIPGENE22473	NPC2	Epididymal secretory protein E1	P79345	ENSCINP00000030926.1	
AIPGENE22527	NPC2	Protein NPC2 homolog	Q9VQ62	ENSCINP00000021760.2	
AIPGENE3074	NPC2	Protein NPC2 homolog	Q9VQ62	ENSCINP00000021760.2	
AIPGENE26780	NPHP3	Nephrocystin-3	Q7Z494	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE4737	NPHP3	Nephrocystin-3	Q7Z494	ENSCINP00000034474.1	
AIPGENE16407	NRF6	Nose resistant to fluoxetine protein 6	Q09225	ENSCINP00000013754.3	GO:0016020 GO:0016021 GO:0016747
AIPGENE16438	NRF6	Nose resistant to fluoxetine protein 6	Q09225	ENSCINP00000013754.3	GO:0016020 GO:0016021 GO:0016747
AIPGENE11352	NWD1	NACHT domain- and WD repeat-containing protein 1	A6H603	ENSCINP00000035575.1	
AIPGENE8824	OBL	Obelin	Q27709		
AIPGENE11507	OCTL	Organic cation transporter-like protein	Q95R48	ENSCINP00000015651.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE17506	OCTL	Organic cation transporter-like protein	Q95R48	ENSCINP00000001843.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE22419	OCTL	Organic cation transporter-like protein	Q95R48	ENSCINP00000020124.3	GO:0005887 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE3952	OCTL	Organic cation transporter-	Q95R48	ENSCINP00000020124.3	GO:0005887

		like protein			GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE25700	OIT3	Oncoprotein-induced transcript 3 protein	Q6V0K7		
AIPGENE6868	OIT3	Oncoprotein-induced transcript 3 protein	Q29RU2		
AIPGENE24712	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000035947.1	GO:0005452 GO:0005887 GO:0015347 GO:0015698 GO:0016020 GO:0016021 GO:0022857 GO:0043252 GO:0055085
AIPGENE5596	ORCT	Organic cation transporter protein	Q9VCA2	ENSCINP00000001843.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE3394	OX1R	Orexin receptor type 1	Q0GBZ5	ENSCINP00000015323.3	
AIPGENE11869	P20D2	Peptidase M20 domain-containing protein 2	Q8IYS1		
AIPGENE18168	P3H2	Prolyl 3-hydroxylase 2	Q8IVL5	ENSCINP00000012097.3	GO:0005506 GO:0016491 GO:0016705 GO:0031418 GO:0055114
AIPGENE5609	P4HA2	Prolyl 4-hydroxylase subunit alpha-2	O15460	ENSCINP00000031242.1	
AIPGENE19666	PA2	Phospholipase A2	A7LCJ2	ENSCINP00000028223.2	
AIPGENE4696	PA2	Phospholipase A2	D2X8K2	ENSCINP00000028223.2	
AIPGENE9806	PA24A	Cytosolic phospholipase A2	P50392	ENSCINP00000017628.3	GO:0004620 GO:0004623 GO:0005737 GO:0006629 GO:0008152 GO:0009395 GO:0016042 GO:0016787 GO:0046872
AIPGENE24257	PAN2	PAB-dependent poly(A)-specific ribonuclease subunit PAN2	Q504Q3	ENSCINP00000000762.3	GO:0000289 GO:0000932 GO:0003676 GO:0004535 GO:0031251 GO:0090503
AIPGENE16667	PAPL	Iron/zinc purple acid phosphatase-like protein	A5D6U8		
AIPGENE16909	PAR14	Poly [ADP-ribose] polymerase 14	Q2EMV9	ENSCINP00000006033.3	GO:0003950 GO:0016740 GO:0016757
AIPGENE22924	PAX3	Paired box protein Pax-3	P24610	ENSCINP00000013350.3	GO:0003677 GO:0005634 GO:0006351 GO:0006355 GO:0007275 GO:0043565
AIPGENE4570	PCFT	Proton-coupled folate transporter	Q6DCX5	ENSCINP00000001721.3	GO:0016020 GO:0016021 GO:0055085
AIPGENE3484	PCP	Lysosomal Pro-X carboxypeptidase	Q5RBU7	ENSCINP00000014308.3	GO:0003085 GO:0004185 GO:0006508 GO:0008236 GO:0008239 GO:0043535 GO:0060055
AIPGENE20867	PCSK5	Proprotein convertase subtilisin/kexin type 5	P41413	ENSCINP00000010394.3	GO:0004252 GO:0005615

					GO:0006508 GO:0008233 GO:0008236 GO:0016485 GO:0016787
AIPGENE1809	PDK1L	Serine/threonine-protein kinase pdik1l	Q32PP3	ENSCINP00000006268.3	GO:0000166 GO:0004672 GO:0004674 GO:0005524 GO:0006468
AIPGENE2323	PEBPH	Phosphatidylethanolamine-binding protein homolog R644	Q5UR88	ENSCINP00000007360.3	
AIPGENE24553	PGDH	15-hydroxyprostaglandin dehydrogenase [NAD(+)]	Q8VCC1	ENSCINP00000002542.2	GO:0016491 GO:0055114
AIPGENE6172	PGRC1	Membrane-associated progesterone receptor component 1	O00264	ENSCINP00000033808.1	
AIPGENE25099	PHLD	Phosphatidylinositol-glycan-specific phospholipase D	Q8R2H5	ENSCINP00000032069.1	
AIPGENE6449	PHNX	Phosphonoacetaldehyde hydrolase	Q2SHM4	ENSCINP00000024234.2	
AIPGENE11495	PHR	Deoxyribodipyrimidine photo-lyase	Q28811		
AIPGENE11501	PISD	Phosphatidylserine decarboxylase proenzyme	Q5R8I8	ENSCINP00000036014.1	GO:0004609 GO:0008654
AIPGENE23141	PK1L2	Polycystic kidney disease protein 1-like 2	Q7TN88	ENSCINP00000008806.3	GO:0005262 GO:0005509 GO:0016020 GO:0016021 GO:0050982 GO:0070588
AIPGENE940	PK1L2	Polycystic kidney disease protein 1-like 2	Q7TN88	ENSCINP00000008806.3	GO:0005262 GO:0005509 GO:0016020 GO:0016021 GO:0050982 GO:0070588
AIPGENE23144	PKDRE	Polycystic kidney disease and receptor for egg jelly-related protein	Q9Z0T6	ENSCINP00000034979.1	GO:0001822 GO:0005262 GO:0016020 GO:0016021 GO:0050982 GO:0070588
AIPGENE17365	PKHL1	Fibrocystin-L	Q80ZA4	ENSCINP00000032083.1	
AIPGENE762	PLBL2	Putative phospholipase B-like 2	Q2KIY5	ENSCINP00000021735.2	
AIPGENE11529	PLD2	Phospholipase D2	O14939	ENSCINP00000026201.2	
AIPGENE27672	PLK4	Serine/threonine-protein kinase PLK4	A7SNN5	ENSCINP00000028801.2	GO:0016020 GO:0016021
AIPGENE1675	POL2	Retrovirus-related Pol polyprotein from transposon 297	P20825	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE19842	POL2	Retrovirus-related Pol polyprotein from transposon 297	P20825	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE27952	POL2	Retrovirus-related Pol polyprotein from transposon 297	P20825	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE5480	POL3	Retrovirus-related Pol polyprotein from transposon 17.6	P04323	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE16289	POL5	Retrovirus-related Pol polyprotein from transposon opus	Q8I7P9	ENSCINP00000036259.1	GO:0003676 GO:0008270 GO:0015074
AIPGENE27985	PP16A	Protein phosphatase 1 regulatory subunit 16A	Q96I34	ENSCINP00000008656.3	GO:0007165 GO:0019208 GO:0019901 GO:0050790
AIPGENE12995	PPIB	Peptidyl-prolyl cis-trans isomerase B	P24367	ENSCINP00000004181.2	GO:0000413 GO:0003755 GO:0006457 GO:0016853
AIPGENE7602	PPT1	Palmitoyl-protein thioesterase 1	Q8HXW6	ENSCINP00000016630.3	GO:0002084 GO:0005764

					GO:0006898 GO:0008474 GO:0098599
AIPGENE3994	PRDM6	Putative histone-lysine N-methyltransferase PRDM6	Q9NQX0	ENSCINP00000006383.3	GO:0003676 GO:0046872
AIPGENE4010	PRDM6	Putative histone-lysine N-methyltransferase PRDM6	Q3UZD5	ENSCINP00000007111.3	GO:0003676 GO:0046872
AIPGENE16440	PRPF3	U4/U6 small nuclear ribonucleoprotein Prp3	Q5ZJ85	ENSCINP00000006800.3	GO:0000398 GO:0046540
AIPGENE1816	PRS23	Serine protease 23	Q9D6X6	ENSCINP00000030817.1	GO:0004252 GO:0006508
AIPGENE27212	PSA	Puromycin-sensitive aminopeptidase	P55786	ENSCINP00000014130.3	GO:0005737 GO:0005886 GO:0006508 GO:0008237 GO:0008270 GO:0042277 GO:0043171 GO:0070006
AIPGENE8998	PXDN	Peroxidasin homolog	Q92626	ENSCINP00000014324.3	GO:0004601 GO:0006979 GO:0020037 GO:0055114 GO:0098869
AIPGENE9002	PXMP2	Peroxisomal membrane protein 2	Q07066	ENSCINP00000026931.2	
AIPGENE1976	QORX	Quinone oxidoreductase PIG3	Q53FA7	ENSCINP00000016387.3	GO:0008270 GO:0016491 GO:0055114
AIPGENE20314	RAB32	Ras-related protein Rab-32	Q9CZE3	ENSCINP00000013519.3	GO:0005525 GO:0005739 GO:0005802 GO:0007264 GO:0032438 GO:0042470
AIPGENE10740	RAB7L	Ras-related protein Rab-7L1	Q63481	ENSCINP00000013519.3	GO:0005525 GO:0005739 GO:0005802 GO:0007264 GO:0032438 GO:0042470
AIPGENE5601	RACA	Rho-related protein racA	P34147	ENSCINP00000035599.1	
AIPGENE15688	RBCC1	RB1-inducible coiled-coil protein 1	Q9ESK9	ENSCINP00000004276.3	GO:0000045 GO:0001934 GO:0019901 GO:0034045 GO:1990316
AIPGENE13372	RBM34	RNA-binding protein 34	P42696	ENSCINP00000035044.1	GO:0000166 GO:0003676
AIPGENE13091	RBM5	RNA-binding protein 5	P52756	ENSCINP00000005260.3	GO:0000166 GO:0003676
AIPGENE334	RCN3	Reticulocalbin-3	Q8BH97	ENSCINP00000025165.2	GO:0005509
AIPGENE16826	RDHE2	Epidermal retinol dehydrogenase 2	Q7TQA3	ENSCINP00000002227.3	GO:0016020 GO:0016021
AIPGENE25745	REG1A	Lithostathine-1-alpha	P05451	ENSCINP00000030082.1	
AIPGENE14277	RERG	Ras-related and estrogen-regulated growth inhibitor	Q0VCJ7	ENSCINP00000026160.2	
AIPGENE2133	RET	Proto-oncogene tyrosine-protein kinase receptor Ret	P35546	ENSCINP00000007167.3	GO:0004672 GO:0004713 GO:0004714 GO:0005524 GO:0006468 GO:0007169 GO:0016020 GO:0016021 GO:0018108
AIPGENE18105	RHBG	Ammonium transporter Rh type B	Q7T070	ENSCINP00000018176.3	GO:0005887 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488

AIPGENE18108	RHCG	Ammonium transporter Rh type C	Q19KH7	ENSCINP00000018176.3	GO:0005887 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488
AIPGENE13241	RHCL2	Ammonium transporter Rh type C-like 2	Q8JI14	ENSCINP00000018176.3	GO:0005887 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488
AIPGENE27642	RHG06	Rho GTPase-activating protein 6	O43182	ENSCINP00000008340.3	GO:0007165 GO:0008289
AIPGENE18190	RN213	E3 ubiquitin-protein ligase RNF213	Q63HN8	ENSCINP00000025063.2	
AIPGENE3139	RNOY	Ribonuclease Oy	Q7M456	ENSCINP00000030837.1	
AIPGENE8557	ROCO4	Probable serine/threonine-protein kinase roco4	Q6XHB2	ENSCINP00000018103.3	
AIPGENE2264	ROMO1	Reactive oxygen species modulator 1	A4QNF3		
AIPGENE2707	RPB1	DNA-directed RNA polymerase II subunit RPB1 (Fragment)	P11414		
AIPGENE1565	RRM3	ATP-dependent DNA helicase RRM3	P38766	ENSCINP00000005776.3	GO:0000002 GO:0000166 GO:0000723 GO:0003677 GO:0003678 GO:0004386 GO:0005524 GO:0005634 GO:0005657 GO:0005739 GO:0006260 GO:0006281 GO:0006310 GO:0006974 GO:0010521 GO:0016787 GO:0032211 GO:0032508 GO:0043141 GO:0044806 GO:0051974
AIPGENE18289	RTJK	RNA-directed DNA polymerase from mobile element jockey	P21328	ENSCINP00000031852.1	
AIPGENE1439	S13A3	Solute carrier family 13 member 3	Q8WWT9	ENSCINP00000017609.3	GO:0005215 GO:0005887 GO:0006814 GO:0016020 GO:0016021 GO:0055085 GO:0098656
AIPGENE1279	SLC15A4	Solute carrier family 15 member 4	Q68F72	ENSCINP00000014849.3	GO:0005215 GO:0006810 GO:0006857 GO:0016020 GO:0016021
AIPGENE27783	SLC15A4	Solute carrier family 15 member 4	Q68F72	ENSCINP00000000708.3	GO:0005215 GO:0006810 GO:0006857 GO:0016020 GO:0016021
AIPGENE27827	SLC15A4	Solute carrier family 15 member 4	O09014	ENSCINP00000033357.1	
AIPGENE20198	S1C2A	Sulfotransferase 1C2A	Q9WUW9	ENSCINP00000035987.1	
AIPGENE5971	S22A5	Solute carrier family 22 member 5	O76082	ENSCINP00000015651.3	GO:0005215 GO:0016020

					GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE20675	S22A9	Solute carrier family 22 member 9	Q8IVM8	ENSCINP00000015651.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE3453	S22AA	Solute carrier family 22 member 10	Q63ZE4	ENSCINP00000006844.3	GO:0005887 GO:0008514 GO:0015711 GO:0016020 GO:0016021 GO:0022857 GO:0055085
AIPGENE3502	S22AG	Solute carrier family 22 member 16	Q66KG0	ENSCINP00000031102.1	GO:0016020 GO:0016021
AIPGENE26569	S238B	Solute carrier family 25 member 38-B	P0CAT2	ENSCINP00000018439.3	
AIPGENE27216	S23A1	Solute carrier family 23 member 1	Q9Z2J0	ENSCINP00000002802.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE24231	S23A2	Solute carrier family 23 member 2	B0JZG0	ENSCINP00000007695.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE8501	S26A4	Pendrin	Q9R154	ENSCINP00000022644.2	
AIPGENE3906	S6A11	Sodium- and chloride-dependent GABA transporter 3	P48066	ENSCINP00000017435.3	
AIPGENE6454	S6A13	Sodium- and chloride-dependent GABA transporter 2	P31646	ENSCINP00000017435.3	
AIPGENE21470	SAP	Proactivator polypeptide	O13035	ENSCINP00000013045.3	GO:0001664 GO:0005737 GO:0005764 GO:0006629 GO:0006665 GO:0008047 GO:0043085
AIPGENE17132	SAS6	Spindle assembly abnormal protein 6 homolog	Q5ZMV2	ENSCINP00000018623.3	
AIPGENE13749	SBP1	Selenium-binding protein 1	Q569D5	ENSCINP00000010226.3	GO:0008430
AIPGENE17806	SC5A3	Sodium/myo-inositol cotransporter	P31637	ENSCINP00000014618.3	
AIPGENE25227	SC5A3	Sodium/myo-inositol cotransporter	P53794	ENSCINP00000031732.1	GO:0007165
AIPGENE14217	SC5A7	High-affinity choline transporter 1	Q8UWF0	ENSCINP00000008776.3	GO:0005215 GO:0005307 GO:0005887 GO:0006810 GO:0008292 GO:0015871 GO:0016020 GO:0016021 GO:0033265 GO:0055085
AIPGENE10021	SC5A8	Sodium-coupled monocarboxylate transporter 1	Q8BYF6	ENSCINP00000011213.3	GO:0005215 GO:0005887 GO:0006810 GO:0015293 GO:0016020 GO:0016021 GO:0055085
AIPGENE23673	SC5D	Lathosterol oxidase	O88822	ENSCINP00000027428.2	
AIPGENE18339	SCAP	Sterol regulatory element-binding protein cleavage-activating protein	Q6GQT6	ENSCINP00000016846.3	GO:0000139 GO:0005789 GO:0008203 GO:0015485

					GO:0016020 GO:0016021 GO:0032933
AIPGENE3555	SCPDL	Saccharopine dehydrogenase-like oxidoreductase	Q8NBX0	ENSCINP00000023327.2	GO:0016020 GO:0016021 GO:0016491 GO:0055114
AIPGENE21060	SCRK	Putative fructokinase	O05510		
AIPGENE24684	SDF2	Stromal cell-derived factor 2	Q3SZ45	ENSCINP00000029849.1	GO:0000032 GO:0004169 GO:0007275 GO:0016020 GO:0031502 GO:0035269 GO:0044845 GO:1900101
AIPGENE23099	SE1L1	Protein sel-1 homolog 1	Q9UBV2	ENSCINP00000006011.3	GO:0016020 GO:0016021
AIPGENE24751	SEC	Probable UDP-N-acetylglucosamine—peptide N-acetylglucosaminyltransferase SEC	Q9M8Y0	ENSCINP00000013052.3	
AIPGENE17639	SEM5B	Semaphorin-5B	Q9P283	ENSCINP00000015858.3	GO:0005509
AIPGENE11035	SFRP3	Secreted frizzled-related protein 3	P97401	ENSCINP00000036216.1	
AIPGENE12684	SGPL1	Sphingosine-1-phosphate lyase 1	Q8CHN6	ENSCINP00000002369.3	GO:0003824 GO:0016829 GO:0016831 GO:0019752 GO:0030170
AIPGENE337	SGS1	ATP-dependent helicase SGS1	P35187	ENSCINP00000008261.3	GO:0003676 GO:0004386 GO:0005524 GO:0006260 GO:0006281 GO:0006310 GO:0008026 GO:0016787 GO:0032508 GO:0043140
AIPGENE8899	SGS4	Salivary glue protein Sgs-4	Q00725		
AIPGENE721	SIAE	Sialate O-acetylesterase	P70665	ENSCINP00000001140.3	
AIPGENE11218	SIR1	NAD-dependent protein deacetylase SRT1	B8ARK7	ENSCINP00000002463.3	GO:0070403
AIPGENE5456	SL9B2	Mitochondrial sodium/hydrogen exchanger 9B2	Q86UD5	ENSCINP00000006194.3	GO:0006810 GO:0006812 GO:0015299 GO:0016020 GO:0016021 GO:0055085 GO:1902600
AIPGENE10650	SNR27	U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein	Q8K194		
AIPGENE17665	SNTG2	Gamma-2-syntrophin	Q925E0	ENSCINP00000026050.2	
AIPGENE19965	SPD2A	SH3 and PX domain-containing protein 2A	Q5TCZ1	ENSCINP00000024481.2	
AIPGENE2530	SPDEF	SAM pointed domain-containing Ets transcription factor	Q9WTP3	ENSCINP00000024831.2	GO:0000981 GO:0003677 GO:0003700 GO:0005634 GO:0006355 GO:0006357 GO:0030154 GO:0043565
AIPGENE14029	SPIT2	Kunitz-type protease inhibitor 2	Q9WU03	ENSCINP00000017737.3	GO:0004867 GO:0010951
AIPGENE25178	SPRC	SPARC	P34714	ENSCINP00000004847.3	GO:0005509 GO:0005578 GO:0007165
AIPGENE16015	SPY2	Protein 99prout homolog 2	Q9PTL2	ENSCINP00000030183.1	GO:0007275 GO:0009966 GO:0016020

AIPGENE11862	STAR9	StAR-related lipid transfer protein 9	Q9P2P6	ENSCINP00000000065.3	GO:0008289
AIPGENE6178	STT3B	Dolichyl-diphosphooligosaccharide—protein glycosyltransferase subunit STT3B	Q8TCJ2	ENSCINP00000003935.3	GO:0004576 GO:0004579 GO:0006486 GO:0008250 GO:0016020 GO:0016021 GO:0018279 GO:0043687
AIPGENE8937	SYG	Glycine—tRNA ligase	P41250	ENSCINP00000018157.4	
AIPGENE21323	SYIC	Isoleucine—tRNA ligase, cytoplasmic	P41252	ENSCINP00000005068.3	GO:0000166 GO:0002161 GO:0004812 GO:0004822 GO:0005524 GO:0005737 GO:0006412 GO:0006418 GO:0006428 GO:0006450 GO:0016874
AIPGENE12993	SYT1	Synaptotagmin-1	Q5R4J5	ENSCINP00000001641.3	GO:0005509 GO:0005544 GO:0005886 GO:0006887 GO:0006906 GO:0016020 GO:0017158 GO:0019905 GO:0030276 GO:0048791 GO:0098793
AIPGENE13021	SYT7	Synaptotagmin-7	O43581	ENSCINP00000001641.3	GO:0005509 GO:0005544 GO:0005886 GO:0006887 GO:0006906 GO:0016020 GO:0017158 GO:0019905 GO:0030276 GO:0048791 GO:0098793
AIPGENE27010	TBX1	T-box transcription factor TBX1	Q8AXX2	ENSCINP00000008846.2	GO:0003677 GO:0003700 GO:0005634 GO:0006351 GO:0006355
AIPGENE600	TCTP	Translationally-controlled tumor protein homolog	Q7QCK2	ENSCINP00000029118.2	
AIPGENE28871	TDA6	Putative vacuolar protein sorting-associated protein TDA6	Q06466		
AIPGENE25165	TGFR2	TGF-beta receptor type-2	P38438	ENSCINP00000000633.3	GO:0000166 GO:0004672 GO:0004674 GO:0004675 GO:0004702 GO:0005524 GO:0006468 GO:0007178 GO:0016020 GO:0016021 GO:0016301 GO:0016310 GO:0016740 GO:0023014 GO:0046872
AIPGENE18026	TGM1	Protein-glutamine gamma-glutamyltransferase K	P23606	ENSCINP00000010148.3	GO:0003810 GO:0018149 GO:0046872
AIPGENE8588	TICN2	Testican-2	Q9ER58	ENSCINP00000006167.3	GO:0005509 GO:0005578

					GO:0007165
AIPGENE8609	TIMP1	Metalloproteinase inhibitor 1	P50122		
AIPGENE2736	TITIN	Titin	A2ASS6	ENSCINP00000015858.3	GO:0005509
AIPGENE7023	TITIN	Titin	A2ASS6	ENSCINP00000013676.3	GO:0000166 GO:0004672 GO:0005524 GO:0005859 GO:0005865 GO:0006468 GO:0006941 GO:0007015 GO:0008307 GO:0016301 GO:0016310 GO:0016740 GO:0030018 GO:0031430 GO:0045214 GO:0051015 GO:0051371 GO:0071688 GO:0097493
AIPGENE11194	TM189	Transmembrane protein 189	Q99LQ7	ENSCINP00000018140.3	GO:0005737 GO:0016020 GO:0016021 GO:0016567 GO:0031625 GO:0061630
AIPGENE3242	TM53B	Transmembrane protein 53-B	Q6DJC8	ENSCINP00000004026.3	GO:0016020 GO:0016021
AIPGENE15175	TMPS3	Transmembrane protease serine 3	Q8K1T0	ENSCINP00000001165.3	GO:0004252 GO:0005044 GO:0005615 GO:0006508 GO:0006898 GO:0008233 GO:0008236 GO:0016020 GO:0016787
AIPGENE27486	TPM	Tropomyosin	Q8WR63	ENSCINP00000004783.3	
AIPGENE203	TPP1	Tripeptidyl-peptidase 1	Q9EQV6		
AIPGENE20835	TRAF3	TNF receptor-associated factor 3	Q60803	ENSCINP00000027890.2	
AIPGENE1290	TRAF4	TNF receptor-associated factor 4	Q61382	ENSCINP00000003738.3	GO:0008270 GO:0046872
AIPGENE1843	TRAF4	TNF receptor-associated factor 4	Q9BUZ4	ENSCINP00000005981.3	
AIPGENE18406	TRET1	Facilitated trehalose transporter Tret1	A9ZSY3	ENSCINP00000014757.3	
AIPGENE3538	TRIP4	Activating signal cointegrator 1	Q9QXN3	ENSCINP00000008619.3	GO:0003713 GO:0005634 GO:0006355 GO:0006366 GO:0008270 GO:0045893
AIPGENE225	TRPC2	Short transient receptor potential channel 2	Q9R283	ENSCINP00000007988.3	GO:0005216 GO:0005262 GO:0006810 GO:0006811 GO:0016020 GO:0016021 GO:0055085 GO:0070588
AIPGENE3734	TRPC2	Short transient receptor potential channel 2 homolog	O62826	ENSCINP00000007988.3	GO:0005216 GO:0005262 GO:0006810 GO:0006811 GO:0016020 GO:0016021 GO:0055085 GO:0070588
AIPGENE4129	TRPC2	Short transient receptor potential channel 2	Q9R283	ENSCINP00000026899.2	
AIPGENE18440	TRPL	Transient-receptor-potential-	P34586	ENSCINP00000007988.3	GO:0005216

		like protein			GO:0005262 GO:0006810 GO:0006811 GO:0016020 GO:0016021 GO:0055085 GO:0070588
AIPGENE10658	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE16596	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE18099	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000018424.3	
AIPGENE24471	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE24487	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE2837	TTC28	Tetratricopeptide repeat protein 28	Q80XJ3	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE5628	TTC28	Tetratricopeptide repeat protein 28	Q80XJ3	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE14019	TWK7	TwiK family of potassium channels protein 7	P34410	ENSCINP00000032062.1	GO:0005267 GO:0005887 GO:0006810 GO:0006811 GO:0016020 GO:0016021 GO:0022841 GO:0030322 GO:0071805
AIPGENE18751	TYR1	Putative tyrosinase-like protein tyr-1	P34269	ENSCINP00000022936.2	
AIPGENE23409	TYRO	Tyrosinase	P55023	ENSCINP00000008838.3	GO:0008152 GO:0016491 GO:0046872 GO:0055114
AIPGENE27733	TYRO	Tyrosinase	P06845	ENSCINP00000008838.3	GO:0008152 GO:0016491 GO:0046872 GO:0055114
AIPGENE19907	U587	UPF0587 protein v1g245604	A7SJ66	ENSCINP00000032370.1	
AIPGENE3226	UBA5	Ubiquitin-like modifier-activating enzyme 5	Q54C02	ENSCINP00000008729.3	GO:0005829 GO:0008641 GO:0061503 GO:0061504 GO:0071566 GO:0071569
AIPGENE20117	UBC23	Probable ubiquitin-conjugating enzyme E2 23	Q9ZVX1	ENSCINP00000023645.2	
AIPGENE8282	UBP8	Ubiquitin carboxyl-terminal hydrolase 8	P40818	ENSCINP00000023650.2	
AIPGENE12152	UROM	Uromodulin	Q862Z3	ENSCINP00000016476.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE8196	UROM	Uromodulin	P07911	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE15989	USOM3	Uncharacterized skeletal organic matrix protein 3 (Fragment)	B8RJM0		
AIPGENE5267	USOM5	Uncharacterized skeletal organic matrix protein 5	B8VIU6		
AIPGENE13361	USOM7	Uncharacterized skeletal organic matrix protein 7	B8WI85		
AIPGENE20926	VGFR2	Vascular endothelial growth factor receptor 2	P35918	ENSCINP00000015860.3	GO:0005509
AIPGENE27144	VIAAT	Vesicular inhibitory amino acid transporter	Q6PF45	ENSCINP00000017689.3	GO:0003333 GO:0015171

					GO:0016020 GO:0016021
AIPGENE6809	VITRN	Vitrin	Q6UXI7	ENSCINP00000031657.1	
AIPGENE8398	VITRN	Vitrin	Q8VHI5	ENSCINP00000001942.3	GO:0005509
AIPGENE17551	VKT1	Kunitz-type protease inhibitor AXPI-I	P81547	ENSCINP00000033091.1	GO:0004867 GO:0010951
AIPGENE18593	VLDLR	Very low-density lipoprotein receptor	P98166	ENSCINP00000031268.1	GO:0005509 GO:0016020 GO:0016021
AIPGENE5447	VP302	Venom protein 302	P0CJ14	ENSCINP00000031459.1	
AIPGENE28851	WDR5B	WD repeat-containing protein 5B	Q5RE95	ENSCINP00000003013.3	
AIPGENE25231	WHITE	Protein white	Q05360	ENSCINP00000031379.1	
AIPGENE21314	WNT5B	Protein Wnt-5b	Q5NVK2	ENSCINP00000000747.3	GO:0005102 GO:0005109 GO:0005576 GO:0005578 GO:0005615 GO:0007275 GO:0016055 GO:0030182 GO:0045165
AIPGENE22302	WTIP	Wilms tumor protein 1-interacting protein homolog	A8DZE6	ENSCINP00000013640.3	GO:0008270 GO:0046872
AIPGENE19802	XDH	Xanthine dehydrogenase	P10351	ENSCINP00000031513.1	GO:0003824 GO:0004854 GO:0005506 GO:0005829 GO:0009055 GO:0009115 GO:0016491 GO:0016614 GO:0016903 GO:0046872 GO:0050660 GO:0051536 GO:0055114
AIPGENE10082	XERD	Integrase/recombinase xerD homolog	O50655		
AIPGENE12154	Y3136	Clavaminate synthase-like protein At3g21360	Q9LIG0		
AIPGENE12159	Y3136	Clavaminate synthase-like protein At3g21360	Q9LIG0		
AIPGENE3623	Y381	Putative ankyrin repeat protein RF_0381	Q4UMH6	ENSCINP00000007941.3	GO:0016020 GO:0016021 GO:0019887 GO:0030165 GO:0045859
AIPGENE10923	Y928	Putative protein methyltransferase MJ0928	Q58338	ENSCINP00000025027.2	GO:0001510 GO:0008168 GO:0009452
AIPGENE4562	YH24	Putative aminopeptidase W07G4.4	Q27245	ENSCINP00000016449.3	GO:0004177 GO:0005622 GO:0005737 GO:0006508 GO:0008235 GO:0019538 GO:0030145
AIPGENE1862	YI31B	Transposon Ty3-I Gag-Pol polyprotein	Q7LHG5	ENSCINP00000003143.3	
AIPGENE15713	YL126	Putative glutamine amidotransferase YLR126C	Q12288		
AIPGENE11737	YL8A	Uncharacterized methyltransferase-like C25B8.10	Q9UTA8	ENSCINP00000017019.3	
AIPGENE4330	YQJG	Glutathionyl-hydroquinone reductase YqjG	P42620	ENSCINP00000006796.3	
AIPGENE26489	YRD6	Uncharacterized protein K02A2.6	Q09575	ENSCINP00000031157.1	
AIPGENE8949	ZC3HF	Zinc finger CCCH domain-containing protein 15	Q8WU90	ENSCINP00000014158.3	GO:0046872
AIPGENE13023	ZCH24	Zinc finger CCHC domain-containing protein 24	B2RVL6	ENSCINP00000010380.3	GO:0003676 GO:0008270
AIPGENE14090	ZCH24	Zinc finger CCHC domain-	Q8N2G6	ENSCINP00000010380.3	GO:0003676

		containing protein 24			GO:0008270
AIPGENE3571	ZHANG	CREB/ATF bZIP transcription factor	Q91ZR3		
AIPGENE22382	ZN665	Zinc finger protein 665	Q5R8X1	ENSCINP00000001925.3	GO:0003676 GO:0046872
AIPGENE17949	ZNT1	Zinc transporter 1	Q9Y6M5	ENSCINP00000022636.2	
AIPGENE22170	ZPP	ZP domain-containing protein	G8HTB6	ENSCINP00000033680.1	GO:0016020 GO:0016021
AIPGENE10171	Predicted protein	Predicted protein	A7RWA9		
AIPGENE10556	Predicted Protein	Predicted Protein			
AIPGENE10722	Predicted Protein	Predicted Protein			
AIPGENE10738	Predicted Protein	Predicted Protein			
AIPGENE10743	Predicted protein	Predicted protein	V4A2K2		
AIPGENE10749	Predicted protein	Predicted protein	A7SA07		
AIPGENE10853	Predicted Protein	Predicted Protein			
AIPGENE10867	Predicted Protein	Predicted Protein			
AIPGENE1087	Predicted protein	Predicted protein	A7RFG4		
AIPGENE1092	Predicted Protein	Predicted Protein			
AIPGENE1126	Predicted Protein	Predicted Protein			
AIPGENE11316	Predicted protein	Predicted protein	A7SUN8		
AIPGENE11363	Predicted protein	Predicted protein	D2VQ64		
AIPGENE11773	Predicted Protein	Predicted Protein			
AIPGENE11891	Predicted protein	Predicted protein	A7SRT5		
AIPGENE12286	Predicted Protein	Predicted Protein			
AIPGENE12333	Predicted Protein	Predicted Protein			
AIPGENE12420	Predicted Protein	Predicted Protein			
AIPGENE12482	Predicted Protein	Predicted Protein			
AIPGENE12565	Predicted protein	Predicted protein	A0A0B2V4Z1		
AIPGENE12722	Predicted Protein	Predicted Protein			
AIPGENE12737	Predicted protein	Predicted protein	A7SA89		
AIPGENE12842	Predicted protein	Predicted protein	A7SP22		
AIPGENE13174	Predicted Protein	Predicted Protein			
AIPGENE13205	Predicted protein	Predicted protein	A7SDJ3		
AIPGENE13212	Predicted Protein	Predicted Protein	KXJ30184.1		
AIPGENE13407	Predicted Protein	Predicted Protein			
AIPGENE13505	Predicted protein	Predicted protein	W4YV05		
AIPGENE13688	Predicted protein	Predicted protein	A7S3A5	ENSCINP00000031518.1	
AIPGENE13726	Predicted Protein	Predicted Protein			
AIPGENE13864	Predicted Protein	Predicted Protein			
AIPGENE13917	Predicted Protein	Predicted Protein			
AIPGENE14010	Predicted Protein	Predicted Protein			
AIPGENE14015	Predicted Protein	Predicted Protein			
AIPGENE14027	Predicted Protein	Predicted Protein			
AIPGENE14046	Predicted Protein	Predicted Protein			
AIPGENE14095	Predicted Protein	Predicted Protein			
AIPGENE14368	Predicted Protein	Predicted Protein			
AIPGENE14420	Predicted Protein	Predicted Protein			
AIPGENE1444	Predicted Protein	Predicted Protein			
AIPGENE14518	Predicted Protein	Predicted Protein			
AIPGENE14579	Predicted protein	Predicted protein	A7RFR8		
AIPGENE14592	Predicted protein	Predicted protein	A7SFI3	ENSCINP00000008667.3	
AIPGENE14740	Predicted protein	Predicted protein	V4AUI6		
AIPGENE14971	Predicted protein	Predicted protein	A7RLZ2		
AIPGENE15114	Predicted Protein	Predicted Protein			
AIPGENE15194	Predicted protein	Predicted protein	A0A0D2WKZ7		
AIPGENE15860	Predicted Protein	Predicted Protein			
AIPGENE15979	Predicted Protein	Predicted Protein			
AIPGENE16131	Predicted Protein	Predicted Protein			
AIPGENE16164	Predicted Protein	Predicted Protein			
AIPGENE16249	Predicted protein	Predicted protein	R7UA65		
AIPGENE16293	Predicted Protein	Predicted Protein			
AIPGENE16299	Predicted protein	Predicted protein	E7EXN5	ENSCINP00000034291.1	
AIPGENE16386	Predicted protein	Predicted protein	B3RVQ7	ENSCINP00000035970.1	GO:0016020 GO:0016021
AIPGENE16397	Predicted protein	Predicted protein	A7RP72		
AIPGENE16543	Predicted Protein	Predicted Protein			
AIPGENE16566	Predicted Protein	Predicted Protein			
AIPGENE16602	Predicted protein	Predicted protein	W5MY83		
AIPGENE16932	Predicted Protein	Predicted Protein			
AIPGENE16938	Predicted Protein	Predicted Protein			
AIPGENE1705	Predicted Protein	Predicted Protein			

AIPGENE17203	Predicted Protein	Predicted Protein			
AIPGENE17481	Predicted Protein	Predicted Protein			
AIPGENE1754	Predicted protein	Predicted protein	A7SG99		
AIPGENE17637	Predicted protein	Predicted protein	Q0ZST9		
AIPGENE17686	Predicted Protein	Predicted Protein			
AIPGENE17762	Predicted protein	Predicted protein	A7RQQ0	ENSCINP00000032414.1	
AIPGENE18002	Predicted Protein	Predicted Protein			
AIPGENE18009	Predicted Protein	Predicted Protein			
AIPGENE18029	Predicted Protein	Predicted Protein			
AIPGENE18042	Predicted Protein	Predicted Protein			
AIPGENE18102	Predicted Protein	Predicted Protein			
AIPGENE18298	Predicted protein	Predicted protein	I1F319		
AIPGENE18519	Predicted protein	Predicted protein	Q668S8		
AIPGENE18557	Predicted protein	Predicted protein	A7SWR1		
AIPGENE18567	Predicted Protein	Predicted Protein			
AIPGENE18767	Predicted Protein	Predicted Protein			
AIPGENE19197	Predicted protein	Predicted protein	C3ZTB7		
AIPGENE19248	Predicted Protein	Predicted Protein			
AIPGENE19267	Predicted Protein	Predicted Protein			
AIPGENE19468	Predicted Protein	Predicted Protein			
AIPGENE19811	Predicted Protein	Predicted Protein			
AIPGENE20175	Predicted Protein	Predicted Protein			
AIPGENE20176	Predicted Protein	Predicted Protein			
AIPGENE2021	Predicted protein	Predicted protein	A7SD04		
AIPGENE20214	Predicted Protein	Predicted Protein			
AIPGENE2036	Predicted protein	Predicted protein	V4ARM9	ENSCINP00000018912.3	
AIPGENE20726	Predicted protein	Predicted protein	C3Z8R0	ENSCINP00000032921.1	GO:0005802 GO:0016020 GO:0016021
AIPGENE20773	Predicted protein	Predicted protein	A7RTC9		
AIPGENE21014	Predicted protein	Predicted protein	A7SIL6		
AIPGENE21365	Predicted protein	Predicted protein	A7SA37	ENSCINP00000012169.3	GO:0004930 GO:0004935 GO:0005887 GO:0007186 GO:0007267 GO:0016020 GO:0016021 GO:0071880
AIPGENE21411	Predicted protein	Predicted protein	A7SGB4		
AIPGENE21849	Predicted protein	Predicted protein	A7S9I2	ENSCINP00000024905.2	
AIPGENE21926	Predicted Protein	Predicted Protein			
AIPGENE22130	Predicted Protein	Predicted Protein			
AIPGENE22416	Predicted protein	Predicted protein	A7S3Z2		
AIPGENE22585	Predicted protein	Predicted protein	A7SR56		
AIPGENE22726	Predicted Protein	Predicted Protein			
AIPGENE22932	Predicted protein	Predicted protein	I1FDK4		
AIPGENE23311	Predicted protein	Predicted protein	A7SGG2		
AIPGENE23321	Predicted protein	Predicted protein	A7RJ46		
AIPGENE23744	Predicted protein	Predicted protein	A7SMG6		
AIPGENE23792	Predicted protein	Predicted protein	A7SVX9		
AIPGENE23942	Predicted Protein	Predicted Protein			
AIPGENE23975	Predicted Protein	Predicted Protein			
AIPGENE24338	Predicted protein	Predicted protein	A7RSY4		
AIPGENE24466	Predicted Protein	Predicted Protein			
AIPGENE24682	Predicted Protein	Predicted Protein			
AIPGENE25157	Predicted protein	Predicted protein	C3XUM9		
AIPGENE25232	Predicted Protein	Predicted Protein			
AIPGENE25329	Predicted Protein	Predicted Protein			
AIPGENE26110	Predicted protein	Predicted protein	A7SC30		
AIPGENE26147	Predicted Protein	Predicted Protein			
AIPGENE26273	Predicted Protein	Predicted Protein			
AIPGENE26285	Predicted protein	Predicted protein	A7SYZ2		
AIPGENE26386	Predicted protein	Predicted protein	W4YMQ5		
AIPGENE26401	Predicted protein	Predicted protein	B3S0V9		
AIPGENE26413	Predicted protein	Predicted protein	A7RP09		
AIPGENE26433	Predicted Protein	Predicted Protein			
AIPGENE26443	Predicted protein	Predicted protein	A7RP09		
AIPGENE26482	Predicted protein	Predicted protein	A7SR48		
AIPGENE26506	Predicted protein	Predicted protein	K1QZ89		
AIPGENE26586	Predicted protein	Predicted protein	A7RLT4		
AIPGENE26594	Predicted protein	Predicted protein	A7S620		
AIPGENE26614	Predicted protein	Predicted protein	A7RSD2		
AIPGENE26659	Predicted Protein	Predicted Protein			

AIPGENE26798	Predicted Protein	Predicted Protein			
AIPGENE27025	Predicted Protein	Predicted Protein			
AIPGENE27038	Predicted protein	Predicted protein	A7RLP7		
AIPGENE27386	Predicted protein	Predicted protein	A7RN96		
AIPGENE27397	Predicted protein	Predicted protein	A7S3J1		
AIPGENE27441	Predicted Protein	Predicted Protein			
AIPGENE27454	Predicted Protein	Predicted Protein			
AIPGENE27869	Predicted protein	Predicted protein	A7SPS5		
AIPGENE27895	Predicted Protein	Predicted Protein			
AIPGENE28308	Predicted protein	Predicted protein	I1G5Y4	ENSCINP00000027071.2	
AIPGENE28477	Predicted Protein	Predicted Protein			
AIPGENE28713	Predicted protein	Predicted protein	V4C0Z0		
AIPGENE28935	Predicted protein	Predicted protein	C3ZIV2		
AIPGENE2904	Predicted Protein	Predicted Protein			
AIPGENE2958	Predicted protein	Predicted protein	A7SA45		
AIPGENE3065	Predicted protein	Predicted protein	I1FEN7		
AIPGENE312	Predicted protein	Predicted protein	C3ZTQ4		
AIPGENE3220	Predicted Protein	Predicted Protein			
AIPGENE3269	Predicted protein	Predicted protein	A7RN39		
AIPGENE3272	Predicted protein	Predicted protein	A7SPQ9		
AIPGENE335	Predicted protein	Predicted protein	A0A084WKC9		
AIPGENE3381	Predicted Protein	Predicted Protein			
AIPGENE3398	Predicted protein	Predicted protein	A7RTL5		
AIPGENE3677	Predicted protein	Predicted protein	A7SH61		
AIPGENE3833	Predicted Protein	Predicted Protein			
AIPGENE3851	Predicted Protein	Predicted Protein			
AIPGENE3932	Predicted protein	Predicted protein	A7SW67		
AIPGENE4137	Predicted Protein	Predicted Protein			
AIPGENE4174	Predicted Protein	Predicted Protein			
AIPGENE4204	Predicted protein	Predicted protein	A7RT43	ENSCINP00000021573.3	GO:0016020 GO:0016021
AIPGENE4210	Predicted Protein	Predicted Protein	KXJ22156.1		GO:0004672 GO:0005524 GO:0006468 GO:0016020 GO:0016021
AIPGENE4213	Predicted Protein	Predicted Protein	KXJ22189.1		
AIPGENE4240	Predicted protein	Predicted protein	A7RS57		
AIPGENE4275	Predicted Protein	Predicted Protein			
AIPGENE4611	Predicted Protein	Predicted Protein	KXJ11150.1		GO:0000166 GO:0004672 GO:0004713 GO:0004715 GO:0005102 GO:0005524 GO:0006468 GO:0007169 GO:0016301 GO:0016310 GO:0016740 GO:0030154 GO:0031234 GO:0038083 GO:0042127 GO:0045087
AIPGENE4678	Predicted Protein	Predicted Protein			
AIPGENE4800	Predicted protein	Predicted protein	K1R495		
AIPGENE4835	Predicted Protein	Predicted Protein			
AIPGENE5210	Predicted Protein	Predicted Protein			
AIPGENE5319	Predicted protein	Predicted protein	A7SA89		
AIPGENE5343	Predicted protein	Predicted protein	A7SA89		
AIPGENE5354	Predicted protein	Predicted protein	A7SMP6		
AIPGENE5409	Predicted protein	Predicted protein	A7SA37		
AIPGENE5643	Predicted Protein	Predicted Protein			
AIPGENE5777	Predicted Protein	Predicted Protein			
AIPGENE6201	Predicted Protein	Predicted Protein			
AIPGENE6301	Predicted Protein	Predicted Protein			
AIPGENE648	Predicted protein	Predicted protein	A7SRP0		
AIPGENE6480	Predicted protein	Predicted protein	A7S4R8		
AIPGENE6761	Predicted Protein	Predicted Protein			
AIPGENE6765	Predicted protein	Predicted protein	A7SXP8		
AIPGENE6817	Predicted protein	Predicted protein	A7SR51		

AIPGENE6831	Predicted protein	Predicted protein	W4XJ24		
AIPGENE6863	Predicted Protein	Predicted Protein			
AIPGENE7096	Predicted protein	Predicted protein	V4AUI6		
AIPGENE7356	Predicted Protein	Predicted Protein			
AIPGENE7545	Predicted Protein	Predicted Protein			
AIPGENE761	Predicted Protein	Predicted Protein			
AIPGENE7613	Predicted Protein	Predicted Protein			
AIPGENE764	Predicted protein	Predicted protein	A7SIQ2		
AIPGENE7703	Predicted protein	Predicted protein	Q580E4		
AIPGENE784	Predicted Protein	Predicted Protein			
AIPGENE7924	Predicted protein	Predicted protein	A7RWA9		
AIPGENE7992	Predicted protein	Predicted protein	A0A0C5GWD8		
AIPGENE8130	Predicted Protein	Predicted Protein			
AIPGENE8218	Predicted protein	Predicted protein	A7RUC8	ENSCINP00000007943.3	GO:0016020 GO:0016021
AIPGENE8310	Predicted Protein	Predicted Protein			
AIPGENE8385	Predicted Protein	Predicted Protein			
AIPGENE8391	Predicted protein	Predicted protein	A7SP95		
AIPGENE8585	Predicted protein	Predicted protein	A7S659		
AIPGENE8631	Predicted Protein	Predicted Protein			
AIPGENE8653	Predicted protein	Predicted protein	A7SZJ7		
AIPGENE8654	Predicted Protein	Predicted Protein			
AIPGENE8674	Predicted protein	Predicted protein	R7RXF9		
AIPGENE9263	Predicted protein	Predicted protein	A7SFP2		
AIPGENE9289	Predicted protein	Predicted protein	A7SFI3	ENSCINP00000008667.3	
AIPGENE9919	Predicted protein	Predicted protein	A7SIY5	ENSCINP00000020132.3	GO:0005886 GO:0007411 GO:0016020 GO:0016021 GO:0046875 GO:0048013
AIPGENE994	Predicted Protein	Predicted Protein			
AIPGENE9982	Predicted protein	Predicted protein	A7SKQ0		

Table S1-3. Enriched GO terms for the host DEGs unique to the apo-symbiotic individuals

E. diaphana geneID	*Gene name*	*Discription*	*UniprotID*	*C. intestinalis geneID*	*GOID*
AIPGENE19170	*ACBC*	*2-epi-5-epi-valiolone synthase*	*Q9ZAE9*	*ENSCINP00000031198.1*	*GO:0005576 GO:0006030 GO:0008061*
AIPGENE13105	*ACES*	*Acetylcholinesterase*	*Q92035*	*ENSCINP00000009596.3*	*GO:0004104 GO:0005615 GO:0016787 GO:0052689*
AIPGENE18473	*ACES*	*Acetylcholinesterase*	*Q92035*	*ENSCINP00000031848.1*	
AIPGENE8252	*ACTC*	*Actin, cytoplasmic*	*Q964E3*	*ENSCINP00000001538.3*	*GO:0000166 GO:0005524*
AIPGENE28991	*ADA23*	*Disintegrin and metalloproteinase domain-containing protein 23*	*O75077*	*ENSCINP00000018188.3*	*GO:0004222 GO:0005578 GO:0006508 GO:0008237 GO:0008270 GO:0031012*
AIPGENE20164	*ADPRM*	*Manganese-dependent ADP-ribose/CDP-alcohol diphosphatase*	*Q7T291*		
AIPGENE5443	*AGRD1*	*Adhesion G-protein coupled receptor D1*	*A6QLU6*	*ENSCINP00000000119.2*	*GO:0004888 GO:0004930 GO:0007166 GO:0007186 GO:0016020 GO:0016021*
AIPGENE1872	*AGRIN*	*Agrin*	*P31696*		
AIPGENE9489	*AHSA1*	*Activator of 90 kDa heat shock protein ATPase homolog 1*	*O95433*		
AIPGENE20746	*ANPRA*	*Atrial natriuretic peptide receptor 1*	*P16066*	*ENSCINP00000002159.3*	*GO:0000166 GO:0004383 GO:0004672 GO:0005524 GO:0005886 GO:0006182 GO:0006468 GO:0007165 GO:0008074 GO:0009190 GO:0016020 GO:0016021 GO:0016829 GO:0016849 GO:0035556*
AIPGENE23774	*ANT1*	*Major antigen*	*P21249*		
AIPGENE20759	*AOC1*	*Amiloride-sensitive amine oxidase [copper-containing]*	*P36633*		
AIPGENE15174	*AQP10*	*Aquaporin-10*	*Q96PS8*	*ENSCINP00000007510.3*	*GO:0005215 GO:0005887 GO:0006810 GO:0006833 GO:0009992 GO:0015250 GO:0015254 GO:0015793 GO:0016020 GO:0016021 GO:0034220*
AIPGENE13552	*ARF4*	*ADP-ribosylation factor 4*	*P51644*	*ENSCINP00000035274.1*	
AIPGENE347	*ASIC1*	*Acid-sensing ion channel 1*	*Q7T1N4*	*ENSCINP00000013699.3*	*GO:0005272 GO:0006810 GO:0006811 GO:0006814 GO:0016020 GO:0016021 GO:0035725*

AIPGENE17868	*AT2A3*	*Sarcoplasmic/endoplasmic reticulum calcium ATPase 3*	*Q9YGL9*	*ENSCINP00000032608.1*	
AIPGENE27621	*AT2B3*	*Plasma membrane calcium-transporting ATPase 3*	*Q16720*	*ENSCINP00000006789.3*	*GO:0000166 GO:0005388 GO:0005524 GO:0005887 GO:0006810 GO:0006811 GO:0006816 GO:0016020 GO:0016021 GO:0016787 GO:0030165 GO:0043231 GO:0046872 GO:0051480 GO:0070588*
AIPGENE4491	*ATK4*	*Kinesin-4*	*O81635*	*ENSCINP00000031156.1*	
AIPGENE2428	*B4GN4*	*N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase 1*	*Q76KP1*	*ENSCINP00000028502.2*	*GO:0008376 GO:0032580*
AIPGENE27791	*B561I*	*Cytochrome b561 and DOMON domain-containing protein At3g61750*	*Q9M363*	*ENSCINP00000017029.3*	
AIPGENE9079	*BARX2*	*Homeobox protein BarH-like 2*	*Q9UMQ3*	*ENSCINP00000024460.2*	*GO:0003677 GO:0005634 GO:0006355 GO:0043565*
AIPGENE19340	*BTBDG*	*BTB/POZ domain-containing protein 16*	*Q95J53*	*ENSCINP00000000856.3*	
AIPGENE13497	*C163A*	*Scavenger receptor cysteine-rich type 1 protein M130*	*Q2VLG6*	*ENSCINP00000013067.3*	*GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114*
AIPGENE17207	*C209D*	*CD209 antigen-like protein D*	*Q91ZW8*	*ENSCINP00000004174.3*	*GO:0016020 GO:0016021*
AIPGENE19720	*CAH2*	*Carbonic anhydrase 2*	*Q8UWA5*	*ENSCINP00000026304.2*	
AIPGENE5302	*CAH2*	*Carbonic anhydrase 2*	*Q8UWA5*	*ENSCINP00000002042.3*	
AIPGENE25923	*CALUB*	*Calumenin-B*	*B5X4E0*	*ENSCINP00000025165.2*	*GO:0005509*
AIPGENE7715	*CAS4*	*Short-chain collagen C4 (Fragment)*	*P18503*	*ENSCINP00000030234.1*	
AIPGENE27988	*CB070*	*UPF0573 protein C2orf70 homolog*	*A4II40*	*ENSCINP00000015826.3*	
AIPGENE20910	*CEAM1*	*Carcinoembryonic antigen-related cell adhesion molecule 1*	*P31809*	*ENSCINP00000015860.3*	*GO:0005509*
AIPGENE15266	*CH10*	*10 kDa heat shock protein, mitochondrial*	*Q5DC69*	*ENSCINP00000029870.1*	
AIPGENE22859	*CKAP5*	*Cytoskeleton-associated protein 5*	*Q14008*	*ENSCINP00000015580.3*	
AIPGENE6519	*CLIP1*	*CAP-Gly domain-containing linker protein 1*	*P30622*	*ENSCINP00000031236.1*	
AIPGENE18703	*CNBD2*	*Cyclic nucleotide-binding domain-containing protein 2*	*Q95JR6*	*ENSCINP00000031008.1*	
AIPGENE7770	*CNG3*	*Cyclic nucleotide-gated channel rod photoreceptor subunit alpha*	*Q90980*	*ENSCINP00000014322.3*	*GO:0005216 GO:0005249 GO:0005887 GO:0006810 GO:0006811*

					GO:0006813 GO:0016020 GO:0016021 GO:0034220 GO:0042391 GO:0055085 GO:0071805
AIPGENE10284	CO6A3	Collagen alpha-3(VI) chain	P15989	ENSCINP00000031657.1	
AIPGENE12342	CO6A3	Collagen alpha-3(VI) chain	P12111	ENSCINP00000031657.1	
AIPGENE10288	CO6A4	Collagen alpha-4(VI) chain	A2AX52	ENSCINP00000031657.1	
AIPGENE6806	CO6A5	Collagen alpha-5(VI) chain	A8TX70	ENSCINP00000031657.1	
AIPGENE12362	CO6A6	Collagen alpha-6(VI) chain	A6NMZ7	ENSCINP00000031657.1	
AIPGENE22178	CO6A6	Collagen alpha-6(VI) chain	A6NMZ7	ENSCINP00000031657.1	
AIPGENE6812	CO6A6	Collagen alpha-6(VI) chain	A6NMZ7	ENSCINP00000035113.1	
AIPGENE15734	COCA1	Collagen alpha-1(XII) chain	P13944	ENSCINP00000032368.1	GO:0070492
AIPGENE26829	COTL1	Coactosin-like protein	Q2HJ57		
AIPGENE4564	COX15	Cytochrome c oxidase assembly protein COX15 homolog	Q8BJ03	ENSCINP00000001332.3	GO:0004129 GO:0005743 GO:0006784 GO:0008535 GO:0016020 GO:0016021 GO:0016627 GO:0016653 GO:0055114 GO:1902600
AIPGENE14884	COXM2	COX assembly mitochondrial protein 2 homolog	Q8K199	ENSCINP00000026936.2	
AIPGENE23936	CP3A4	Cytochrome P450 3A4	P08684	ENSCINP00000035581.1	GO:0004497 GO:0005506 GO:0016020 GO:0016021 GO:0016491 GO:0016705 GO:0020037 GO:0046872 GO:0055114
AIPGENE27959	CR3L1	Cyclic AMP-responsive element-binding protein 3-like protein 1	Q96BA8	ENSCINP00000011026.3	GO:0003700 GO:0006355 GO:0043565
AIPGENE23405	CR3L2	Cyclic AMP-responsive element-binding protein 3-like protein 2	A1L224	ENSCINP00000011026.3	GO:0003700 GO:0006355 GO:0043565
AIPGENE25182	CRA1B	Collagen alpha-1(XXVII) chain B	A0MSJ1	ENSCINP00000016017.3	
AIPGENE10590	CRIP1	Cysteine-rich protein 1	P50238	ENSCINP00000033008.1	GO:0008270 GO:0046872
AIPGENE7094	CRY1	Cryptochrome-1	Q16526		
AIPGENE22361	CSD	Probable cysteine desulfurase	O27442		
AIPGENE1345	CSGA	C-factor	P21158	ENSCINP00000010325.3	GO:0016491 GO:0055114
AIPGENE7912	CSMD3	CUB and sushi domain-containing protein 3	Q7Z407	ENSCINP00000006993.3	
AIPGENE8956	CSPG2	Versican core protein (Fragments)	Q9ERB4	ENSCINP00000006010.3	GO:0005509 GO:0005886 GO:0007155 GO:0007156 GO:0016020 GO:0016021

AIPGENE28323	*CSRP2*	*Cysteine and glycine-rich protein 2*	*Q05158*	*ENSCINP00000035257.1*	*GO:0008270 GO:0046872*
AIPGENE4505	*CSTF3*	*Cleavage stimulation factor subunit 3*	*Q5RDW9*	*ENSCINP00000013211.3*	*GO:0003729 GO:0005634 GO:0006379 GO:0006396 GO:0006397 GO:0031123*
AIPGENE24288	*CTRL*	*Chymotrypsin-like protease CTRL-1*	*P40313*	*ENSCINP00000013976.2*	*GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787*
AIPGENE122	*CUBN*	*Cubilin*	*O70244*	*ENSCINP00000007799.3*	*GO:0004222 GO:0005509 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787 GO:0046872*
AIPGENE5546	*CY24B*	*Cytochrome b-245 heavy chain*	*Q61093*	*ENSCINP00000008272.3*	*GO:0016020 GO:0016021 GO:0016491 GO:0055114*
AIPGENE4733	*DBP*	*D site-binding protein*	*P16443*	*ENSCINP00000004693.3*	*GO:0003700 GO:0006355 GO:0043565*
AIPGENE28483	*DCA11*	*DDB1- and CUL4-associated factor 11*	*Q5E9I8*	*ENSCINP00000013358.3*	*GO:0005682 GO:0008380 GO:0071011 GO:0071013*
AIPGENE15570	*DCAF8*	*DDB1- and CUL4-associated factor 8*	*Q28I90*	*ENSCINP00000008889.3*	
AIPGENE17925	*DD3*	*Protein DD3-3*	*Q58A42*	*ENSCINP00000028120.2*	
AIPGENE1051	*DEAF1*	*Deformed epidermal autoregulatory factor 1*	*Q24180*	*ENSCINP00000015236.3*	*GO:0003677 GO:0003700 GO:0005634 GO:0006357 GO:0046872 GO:2000026*
AIPGENE14367	*DMBT1*	*Deleted in malignant brain tumors 1 protein*	*Q9UGM3*	*ENSCINP00000013067.3*	*GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114*
AIPGENE22679	*DMD*	*Dystrophin*	*P11533*		
AIPGENE22701	*DMD*	*Dystrophin*	*P11531*	*ENSCINP00000018184.3*	
AIPGENE8405	*DNPEP*	*Aspartyl aminopeptidase*	*Q5RBT2*	*ENSCINP00000030453.1*	*GO:0004177 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787 GO:0046872*
AIPGENE17437	*DOX2*	*Alpha-dioxygenase 2*	*Q9C9U3*	*ENSCINP00000009017.3*	*GO:0004601 GO:0005509 GO:0006979 GO:0016020 GO:0016021 GO:0020037 GO:0055114 GO:0098869*
AIPGENE15774	*DPS1*	*Decaprenyl-diphosphate synthase subunit 1*	*Q5T2R2*	*ENSCINP00000002476.3*	*GO:0008299 GO:0016740*
AIPGENE12166	*DRG1*	*Developmentally-regulated GTP-binding protein 1*	*P43690*	*ENSCINP00000006782.3*	*GO:0005525*
AIPGENE28094	*E2F3*	*Transcription factor E2F3*	*O35261*	*ENSCINP00000030840.1*	*GO:0003677 GO:0003700 GO:0005634*

					GO:0005667 GO:0006351 GO:0006355
AIPGENE19051	ECE1	Endothelin-converting enzyme 1	P42893	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE6882	ECE1	Endothelin-converting enzyme 1	P42893	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE13723	ECE2	Endothelin-converting enzyme 2	Q80Z60	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE18973	ECE2	Endothelin-converting enzyme 2	O60344	ENSCINP00000031762.1	GO:0004222 GO:0006508 GO:0008237
AIPGENE4699	ECI2	Enoyl-CoA delta isomerase 2, mitochondrial	Q5XIC0	ENSCINP00000011768.3	GO:0000062 GO:0003824 GO:0008152
AIPGENE29237	EDIL3	EGF-like repeat and discoidin I-like domain-containing protein 3	O35474	ENSCINP00000006532.3	GO:0005509
AIPGENE21355	ELAV3	ELAV-like protein 3	Q91584	ENSCINP00000009765.3	GO:0000166 GO:0003676 GO:0003723
AIPGENE2247	ERGI1	Endoplasmic reticulum-Golgi intermediate compartment protein 1	Q4V8Y6	ENSCINP00000008339.3	GO:0016020 GO:0016021
AIPGENE9786	EZRI	Ezrin	P31976	ENSCINP00000025580.2	GO:0003779 GO:0005737 GO:0005856 GO:0008092 GO:0019898
AIPGENE7227	FA11	Coagulation factor XI	Q91Y47	ENSCINP00000007068.3	GO:0004252 GO:0005509 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE2779	FA5V	Venom prothrombin activator pseutarin-C non-catalytic subunit	Q7SZN0	ENSCINP00000019787.3	
AIPGENE26162	FAD5	Delta(5) fatty acid desaturase	O74212		
AIPGENE27752	FAT4	Protocadherin Fat 4	Q2PZL6	ENSCINP00000006010.3	GO:0005509 GO:0005886 GO:0007155 GO:0007156 GO:0016020 GO:0016021
AIPGENE5865	FBP1	Fibropellin-1	P10079	ENSCINP00000001293.2	GO:0005509
AIPGENE11620	FOLH1	Glutamate carboxypeptidase 2	O35409	ENSCINP00000014713.3	
AIPGENE14286	FOLH1	Glutamate carboxypeptidase 2	O35409	ENSCINP00000011038.3	GO:0016020 GO:0016021
AIPGENE14305	FOLH1	Glutamate carboxypeptidase 2	O77564	ENSCINP00000014713.3	
AIPGENE27957	FOLH1	Glutamate carboxypeptidase 2	O35409	ENSCINP00000011038.3	GO:0016020 GO:0016021
AIPGENE11097	FUCL	Fucolectin	Q7SIC1		
AIPGENE24084	FXRD2	FAD-dependent oxidoreductase domain-containing protein 2	Q8IWF2	ENSCINP00000013723.3	GO:0004497 GO:0005788 GO:0016491 GO:0030433 GO:0055114
AIPGENE4017	FZD7B	Frizzled-7-B	Q8AVJ9	ENSCINP00000016965.3	GO:0007166 GO:0007275 GO:0016020 GO:0016021 GO:0016055

					GO:0042813 GO:0060070
AIPGENE1086	GABR2	Gamma-aminobutyric acid type B receptor subunit 2	O75899	ENSCINP00000006905.3	GO:0004930 GO:0004965 GO:0007186 GO:0016020 GO:0016021
AIPGENE7219	GABR2	Gamma-aminobutyric acid type B receptor subunit 2	O88871	ENSCINP00000006905.3	GO:0004930 GO:0004965 GO:0007186 GO:0016020 GO:0016021
AIPGENE14619	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000010145.3	GO:0005576
AIPGENE1513	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9H4G4	ENSCINP00000001814.3	GO:0005576
AIPGENE15674	GCDH	Glutaryl-CoA dehydrogenase, mitochondrial	Q2KHZ9	ENSCINP00000013343.3	GO:0000062 GO:0003995 GO:0005739 GO:0008152 GO:0009055 GO:0016491 GO:0016627 GO:0033539 GO:0050660 GO:0052890 GO:0055088 GO:0055114
AIPGENE10496	GCH1	GTP cyclohydrolase 1	Q19980	ENSCINP00000008998.3	GO:0003934 GO:0005525 GO:0005737 GO:0006729 GO:0008270 GO:0046654
AIPGENE29213	GELS1	Gelsolin-like protein 1	Q7JQD3	ENSCINP00000003800.3	GO:0003779 GO:0007010
AIPGENE29236	GELS1	Gelsolin-like protein 1	Q7JQD3	ENSCINP00000022359.2	
AIPGENE8881	GGLO	L-gulonolactone oxidase	Q90YK3	ENSCINP00000035375.1	GO:0003885 GO:0016020 GO:0055114
AIPGENE25057	GLBL3	Beta-galactosidase-1-like protein 3	Q8NCI6	ENSCINP00000024468.2	GO:0004553 GO:0004565 GO:0005773 GO:0005975 GO:0008152 GO:0016787 GO:0016798
AIPGENE18788	GLTP	Glycolipid transfer protein	B0YN54	ENSCINP00000002896.3	GO:0005737 GO:0017089 GO:0046836 GO:0051861
AIPGENE12935	GP161	G-protein coupled receptor 161	B2RPY5	ENSCINP00000019971.3	GO:0004871 GO:0004930 GO:0005887 GO:0007165 GO:0007166 GO:0007186 GO:0016020 GO:0016021
AIPGENE10925	GPR98	G-protein coupled receptor 98	Q6JAN0	ENSCINP00000012945.3	GO:0005432 GO:0006816 GO:0007154 GO:0016020 GO:0016021 GO:0035725 GO:0055085
AIPGENE11030	GPR98	G-protein coupled receptor 98	Q8VHN7	ENSCINP00000011103.2	
AIPGENE24677	GRIA2	Glutamate receptor 2	P19491	ENSCINP00000017550.3	

AIPGENE6351	*GSTM5*	*Glutathione S-transferase Mu 5*	*Q9Z1B2*	*ENSCINP00000022076.2*	*GO:0004364 GO:0008152*
AIPGENE12054	*GTR1*	*Glucose transporter type 1*	*Q8IRI6*	*ENSCINP00000003472.3*	*GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085*
AIPGENE6136	*H3*	*Histone H3*	*P84239*	*ENSCINP00000005615.3*	*GO:0000786 GO:0003677 GO:0005634 GO:0005694 GO:0046982*
AIPGENE26271	*HES1A*	*Transcription factor HES-1-A*	*Q6IRB2*	*ENSCINP00000010322.3*	*GO:0003677 GO:0005634 GO:0006351 GO:0006355 GO:0046983*
AIPGENE11333	*HLF*	*Hepatic leukemia factor*	*Q16534*	*ENSCINP00000016562.3*	*GO:0000122 GO:0001078 GO:0003700 GO:0005634 GO:0006355 GO:0007623 GO:0043565*
AIPGENE11126	*HMCN1*	*Hemicentin-1*	*Q96RW7*	*ENSCINP00000015858.3*	*GO:0005509*
AIPGENE19491	*HMCN1*	*Hemicentin-1*	*Q96RW7*	*ENSCINP00000015860.3*	*GO:0005509*
AIPGENE25044	*HMCN2*	*Hemicentin-2*	*Q8NDA2*	*ENSCINP00000015858.3*	*GO:0005509*
AIPGENE20646	*HMU*	*Halomucin*	*Q18DN4*		
AIPGENE20749	*ICE2*	*Little elongation complex subunit 2*	*Q659A1*		
AIPGENE8006	*IF4E*	*Eukaryotic translation initiation factor 4E*	*P63074*	*ENSCINP00000008672.3*	*GO:0003723 GO:0003743 GO:0005737 GO:0006412 GO:0006413*
AIPGENE27755	*IMSP1*	*Insoluble matrix shell protein 1 (Fragment)*	*P86982*		
AIPGENE10893	*IQCH*	*IQ domain-containing protein H*	*Q9D2K4*	*ENSCINP00000006749.3*	
AIPGENE10972	*K1161*	*Uncharacterized family 31 glucosidase KIAA1161*	*Q69ZQ1*	*ENSCINP00000032199.1*	*GO:0003824 GO:0004553 GO:0005975 GO:0008152 GO:0016787 GO:0016798*
AIPGENE1479	*KCNA1*	*Potassium voltage-gated channel subfamily A member 1*	*Q09470*	*ENSCINP00000023220.2*	
AIPGENE25654	*KCNK9*	*Potassium channel subfamily K member 9*	*Q63ZI0*	*ENSCINP00000032062.1*	*GO:0005267 GO:0005887 GO:0006810 GO:0006811 GO:0016020 GO:0016021 GO:0022841 GO:0030322 GO:0071805*
AIPGENE18117	*KDM4A*	*Lysine-specific demethylase 4A*	*Q8BW72*	*ENSCINP00000004028.3*	*GO:0005634 GO:0008270 GO:0046872*
AIPGENE1019	*KIF27*	*Kinesin-like protein KIF27*	*Q7M6Z4*		
AIPGENE1057	*KIF27*	*Kinesin-like protein KIF27*	*Q7M6Z5*	*ENSCINP00000023041.2*	*GO:0000166 GO:0003777 GO:0005524 GO:0005874 GO:0007018 GO:0008017*

AIPGENE28503	*KIF7*	*Kinesin-like protein kif7*	*Q58G59*		
AIPGENE4464	*L12R1*	*Loss of heterozygosity 12 chromosomal region 1 protein*	*Q969J3*		
AIPGENE530	*LIPR2*	*Pancreatic lipase-related protein 2*	*P81139*	*ENSCINP00000014633.3*	*GO:0004806 GO:0005576 GO:0006629 GO:0052689*
AIPGENE10774	*LONF3*	*LON peptidase N-terminal domain and RING finger protein 3*	*Q9D4H7*	*ENSCINP00000026615.2*	
AIPGENE23304	*LONF3*	*LON peptidase N-terminal domain and RING finger protein 3*	*Q9D4H7*	*ENSCINP00000026615.2*	
AIPGENE11163	*LONP2*	*Lon protease homolog 2, peroxisomal*	*Q5PQY6*	*ENSCINP00000018649.3*	*GO:0000166 GO:0003697 GO:0004176 GO:0004252 GO:0005524 GO:0005759 GO:0006508 GO:0006515 GO:0007005 GO:0008233 GO:0008236 GO:0016787 GO:0016887 GO:0030163 GO:0034599 GO:0043565 GO:0051131 GO:0070361 GO:0070407 GO:0090296*
AIPGENE25932	*LPHN3*	*Latrophilin-3*	*Q9HAR2*	*ENSCINP00000011539.3*	*GO:0004871 GO:0004888 GO:0004930 GO:0005509 GO:0005886 GO:0007155 GO:0007156 GO:0007165 GO:0007166 GO:0007186 GO:0016020 GO:0016021*
AIPGENE29099	*LRRC9*	*Leucine-rich repeat-containing protein 9*	*A0JM56*	*ENSCINP00000000977.3*	
AIPGENE27103	*LRWD1*	*Leucine-rich repeat and WD repeat-containing protein 1*	*B0JZ65*	*ENSCINP00000034343.1*	
AIPGENE24766	*LTXA*	*Leukotoxin*	*P16462*		
AIPGENE23317	*MASY*	*Malate synthase*	*P95329*		
AIPGENE868	*MATN1*	*Cartilage matrix protein*	*P05099*	*ENSCINP00000026363.2*	
AIPGENE2457	*MCFD2*	*Multiple coagulation factor deficiency protein 2 homolog*	*Q8K5B3*	*ENSCINP00000034523.1*	
AIPGENE1364	*MEP1B*	*Meprin A subunit beta*	*Q16820*	*ENSCINP00000017083.3*	*GO:0004222 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016020 GO:0016787 GO:0046872*
AIPGENE20636	*METK2*	*S-adenosylmethionine*	*Q3THS6*	*ENSCINP00000016717.3*	*GO:0000166 GO:0004478*

		synthase isoform type-2			*GO:0005524 GO:0005829 GO:0006556 GO:0006730 GO:0016740 GO:0046872*
AIPGENE6603	*MFGM*	*Lactadherin*	*P70490*	*ENSCINP00000019787.3*	
AIPGENE13503	*MLP*	*Mucin-like protein (Fragment)*	*B3EWY9*	*ENSCINP00000010385.3*	*GO:0005201 GO:0005509 GO:0005578*
AIPGENE21137	*MLP*	*Mucin-like protein (Fragment)*	*B3EWY9*	*ENSCINP00000029354.2*	
AIPGENE28803	*MLP*	*Mucin-like protein (Fragment)*	*B3EWY9*	*ENSCINP00000001964.3*	*GO:0005509*
AIPGENE28887	*MLP*	*Mucin-like protein (Fragment)*	*B3EWY9*	*ENSCINP00000029354.2*	
AIPGENE5096	*MLP*	*Mucin-like protein (Fragment)*	*B3EWY9*	*ENSCINP00000026365.2*	
AIPGENE11485	*MLRP1*	*MAM and LDL-receptor class A domain-containing protein 1 (Fragment)*	*B3EWZ5*	*ENSCINP00000036530.1*	*GO:0016020*
AIPGENE11469	*MLRP2*	*MAM and LDL-receptor class A domain-containing protein 2 (Fragment)*	*B3EWZ6*	*ENSCINP00000033374.1*	*GO:0016020*
AIPGENE21266	*MLRP2*	*MAM and LDL-receptor class A domain-containing protein 2 (Fragment)*	*B3EWZ6*	*ENSCINP00000024995.2*	*GO:0016020*
AIPGENE22528	*MMGT1*	*Membrane magnesium transporter 1*	*Q8N4V1*		
AIPGENE2967	*MOXD1*	*DBH-like monooxygenase protein 1 homolog*	*Q5TZ24*	*ENSCINP00000032889.1*	
AIPGENE15096	*MOXD2*	*DBH-like monooxygenase protein 2 homolog*	*Q08CS6*	*ENSCINP00000005555.3*	*GO:0003824 GO:0004497 GO:0005507 GO:0016715 GO:0055114*
AIPGENE331	*MRC1*	*Macrophage mannose receptor 1*	*P22897*	*ENSCINP00000004174.3*	*GO:0016020 GO:0016021*
AIPGENE14071	*MUC5B*	*Mucin-5B*	*Q9HC84*	*ENSCINP00000035305.1*	
AIPGENE4552	*MUC5B*	*Mucin-5B*	*Q9HC84*	*ENSCINP00000035305.1*	
AIPGENE8390	*MUP4*	*Transmembrane matrix receptor MUP-4*	*Q21281*	*ENSCINP00000026434.2*	
AIPGENE16461	*MYL6B*	*Myosin light chain 6B*	*P14649*	*ENSCINP00000009959.3*	*GO:0005509*
AIPGENE23854	*MYO1D*	*Unconventional myosin-Id*	*Q63357*	*ENSCINP00000018059.3*	*GO:0000166 GO:0003774 GO:0003779 GO:0005524 GO:0016459*
AIPGENE8265	*MYS*	*Myosin heavy chain, striated muscle*	*P24733*	*ENSCINP00000036000.1*	
AIPGENE18853	*NID1*	*Nidogen-1*	*P14543*	*ENSCINP00000010385.3*	*GO:0005201 GO:0005509 GO:0005578*
AIPGENE26957	*NID1*	*Nidogen-1*	*P10493*	*ENSCINP00000024948.2*	*GO:0005509 GO:0016020 GO:0016021*
AIPGENE18468	*NLGNX*	*Neuroligin-4, X-linked*	*Q8N0W4*	*ENSCINP00000017593.3*	*GO:0005615 GO:0016787 GO:0052689*
AIPGENE18542	*NLGNX*	*Neuroligin-4, X-linked*	*Q8N0W4*	*ENSCINP00000034802.1*	*GO:0005615 GO:0016787 GO:0052689*

AIPGENE13702	*NNMT*	*Nicotinamide N-methyltransferase*	*O55239*		
AIPGENE3676	*NOGG*	*Noggin*	*P97466*	*ENSCINP00000002170.2*	*GO:0001501 GO:0001649 GO:0005615 GO:0007399 GO:0030514 GO:0045596*
AIPGENE1433	*NRP2*	*Neuropilin-2*	*O35375*	*ENSCINP00000019787.3*	
AIPGENE10916	*NT5D3*	*5'-nucleotidase domain-containing protein 3*	*Q86UY8*	*ENSCINP00000017103.3*	*GO:0008253 GO:0016311 GO:0046872*
AIPGENE21041	*NUD14*	*Uridine diphosphate glucose pyrophosphatase*	*Q05B60*	*ENSCINP00000022311.2*	*GO:0016787 GO:0016818 GO:0046872*
AIPGENE12702	*OPSD*	*Rhodopsin*	*O93459*	*ENSCINP00000000739.3*	*GO:0004871 GO:0004930 GO:0005887 GO:0007165 GO:0007186 GO:0007187 GO:0007218 GO:0007268 GO:0016020 GO:0016021 GO:0042923 GO:0043005*
AIPGENE24046	*OPSD1*	*Rhodopsin, GQ-coupled*	*O15973*	*ENSCINP00000030239.1*	
AIPGENE11859	*ORCT*	*Organic cation transporter protein*	*Q9VCA2*	*ENSCINP00000015651.3*	*GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085*
AIPGENE17500	*ORCT*	*Organic cation transporter protein*	*Q9VCA2*	*ENSCINP00000015651.3*	*GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085*
AIPGENE19791	*ORCT*	*Organic cation transporter protein*	*Q9VCA2*	*ENSCINP00000027190.2*	
AIPGENE26154	*ORCT*	*Organic cation transporter protein*	*Q9VCA2*	*ENSCINP00000020124.3*	*GO:0005887 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085*
AIPGENE22267	*PAO*	*Polyamine oxidase*	*O64411*	*ENSCINP00000021792.2*	*GO:0016491 GO:0055114*
AIPGENE20939	*PCFT*	*Proton-coupled folate transporter*	*Q6DCX5*	*ENSCINP00000001721.3*	*GO:0016020 GO:0016021 GO:0055085*
AIPGENE4302	*PDPK1*	*3-phosphoinositide-dependent protein kinase 1*	*O55173*	*ENSCINP00000014529.3*	
AIPGENE954	*PKDRE*	*Polycystic kidney disease and receptor for egg jelly-related protein*	*Q9NTG1*	*ENSCINP00000035440.1*	*GO:0001822 GO:0005262 GO:0016020 GO:0016021 GO:0050982 GO:0070588*
AIPGENE1824	*PNPO*	*Pyridoxine-5'-phosphate oxidase*	*Q5E9K3*	*ENSCINP00000011903.3*	*GO:0004733 GO:0008615 GO:0010181 GO:0016491 GO:0016638 GO:0042823 GO:0055114*
AIPGENE25457	*POL3*	*Retrovirus-related Pol polyprotein from transposon 17.6*	*P04323*	*ENSCINP00000036259.1*	*GO:0003676 GO:0008270 GO:0015074*

AIPGENE18031	*PPR32*	*Protein phosphatase 1 regulatory subunit 32*	*Q2T9T0*	*ENSCINP00000020145.3*	
AIPGENE17583	*PROP*	*Properdin*	*P27918*	*ENSCINP00000015672.3*	*GO:0005509 GO:0005576 GO:0007155*
AIPGENE19952	*PROSC*	*Proline synthase co-transcribed bacterial homolog protein*	*O94903*	*ENSCINP00000004257.3*	*GO:0005622 GO:0030170*
AIPGENE12138	*PRSS8*	*Prostasin*	*Q16651*	*ENSCINP00000002671.3*	*GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787*
AIPGENE19805	*PTHD3*	*Patched domain-containing protein 3*	*Q0EEE2*	*ENSCINP00000008700.3*	*GO:0005319 GO:0006869 GO:0016020 GO:0016021*
AIPGENE6151	*PTPRA*	*Receptor-type tyrosine-protein phosphatase alpha*	*P18052*	*ENSCINP00000030681.1*	*GO:0004721 GO:0004725 GO:0006470 GO:0016020 GO:0016021 GO:0016311 GO:0016787 GO:0016791 GO:0035335*
AIPGENE23541	*PTPRF*	*Receptor-type tyrosine-protein phosphatase F*	*A4IFW2*	*ENSCINP00000015322.3*	
AIPGENE1429	*PTPRQ*	*Phosphatidylinositol phosphatase PTPRQ*	*O88488*	*ENSCINP00000015322.3*	
AIPGENE21680	*PTPRQ*	*Phosphatidylinositol phosphatase PTPRQ*	*O88488*	*ENSCINP00000015322.3*	
AIPGENE25229	*PTPRQ*	*Phosphatidylinositol phosphatase PTPRQ*	*P0C5E4*	*ENSCINP00000015322.3*	
AIPGENE16036	*PUF60*	*Poly(U)-binding-splicing factor PUF60*	*Q9WV25*	*ENSCINP00000017200.3*	
AIPGENE9739	*QRFPR*	*Pyroglutamylated Rfamide peptide receptor*	*Q96P65*	*ENSCINP00000000421.3*	*GO:0004871 GO:0004930 GO:0005887 GO:0007165 GO:0007186 GO:0007200 GO:0007218 GO:0008528 GO:0016020 GO:0016021*
AIPGENE1366	*RAN*	*GTP-binding nuclear protein Ran*	*P52301*	*ENSCINP00000008166.3*	*GO:0000054 GO:0000166 GO:0003924 GO:0005525 GO:0005634 GO:0005737 GO:0006606 GO:0006810 GO:0006886 GO:0006913 GO:0007165 GO:0007264 GO:0015031*
AIPGENE7612	*RBP2A*	*Reticulocyte-binding protein 2 homolog a*	*Q8IDX6*	*ENSCINP00000013676.3*	*GO:0000166 GO:0004672 GO:0005524 GO:0005859 GO:0005865 GO:0006468 GO:0006941*

					GO:0007015 GO:0008307 GO:0016301 GO:0016310 GO:0016740 GO:0030018 GO:0031430 GO:0045214 GO:0051015 GO:0051371 GO:0071688 GO:0097493
AIPGENE26089	*RNF17*	*RING finger protein 17*	*Q9BXT8*	*ENSCINP00000030398.1*	
AIPGENE26150	*ROR2*	*Tyrosine-protein kinase transmembrane receptor ROR2*	*Q01974*	*ENSCINP00000025646.2*	
AIPGENE20974	*ROST*	*Protein rolling stone*	*O44252*	*ENSCINP00000029351.2*	
AIPGENE11062	*RQL4A*	*ATP-dependent DNA helicase Q-like 4A*	*Q8L840*	*ENSCINP00000004562.3*	*GO:0000166 GO:0000722 GO:0000731 GO:0000784 GO:0003676 GO:0003677 GO:0003824 GO:0004386 GO:0005524 GO:0005622 GO:0005654 GO:0005737 GO:0006260 GO:0006281 GO:0006284 GO:0006302 GO:0006310 GO:0007569 GO:0008026 GO:0009378 GO:0016787 GO:0031297 GO:0032508 GO:0043140 GO:0044237 GO:0071480*
AIPGENE6511	*RTXE*	*Probable RNA-directed DNA polymerase from transposon X-element*	*Q9NBX4*		
AIPGENE5684	*RU17*	*U1 small nuclear ribonucleoprotein 70 kDa*	*Q55FQ0*	*ENSCINP00000012532.3*	*GO:0000166 GO:0000243 GO:0000398 GO:0003676 GO:0003729 GO:0005685 GO:0030619 GO:0071004 GO:0071011*
AIPGENE20174	*S18B1*	*MFS-type transporter SLC18B1*	*D3Z5L6*	*ENSCINP00000030034.1*	*GO:0016020 GO:0016021 GO:0055085*
AIPGENE1245	*SAM15*	*Sterile alpha motif domain-containing protein 15*	*Q95JY5*	*ENSCINP00000005696.3*	
AIPGENE789	*SAP3*	*Ganglioside GM2 activator*	*Q60648*		
AIPGENE26536	*SC5A7*	*High-affinity choline transporter 1*	*O02228*	*ENSCINP00000008788.3*	*GO:0005215 GO:0005307 GO:0005887 GO:0006810 GO:0008292 GO:0015871*

					GO:0016020 GO:0016021 GO:0033265 GO:0055085
AIPGENE6458	SC6A1	Sodium- and chloride-dependent GABA transporter 1	P30531	ENSCINP00000014199.3	GO:0005328 GO:0005887 GO:0006810 GO:0006836 GO:0015293 GO:0015375 GO:0016020 GO:0016021 GO:0036233 GO:0055085 GO:0060012
AIPGENE9486	SCN4A	Sodium channel protein type 4 subunit alpha	P15390	ENSCINP00000004967.3	GO:0001518 GO:0005216 GO:0005244 GO:0005248 GO:0005272 GO:0005886 GO:0006810 GO:0006811 GO:0006814 GO:0016020 GO:0016021 GO:0019228 GO:0034220 GO:0034765 GO:0035725 GO:0055085 GO:0060078 GO:0086010
AIPGENE1052	SCO1	Protein SCO1 homolog, mitochondrial	A1A4J8	ENSCINP00000018444.3	
AIPGENE5852	SF3B1	Splicing factor 3B subunit 1	O57683	ENSCINP00000005223.4	GO:0000245 GO:0003729 GO:0005686 GO:0005689 GO:0071004 GO:0071013
AIPGENE9153	SFSWA	Splicing factor, suppressor of white-apricot homolog	Q3USH5	ENSCINP00000025405.2	
AIPGENE2150	SGPP2	Sphingosine-1-phosphate phosphatase 2	Q8IWX5	ENSCINP00000007659.3	GO:0016020 GO:0016021
AIPGENE20701	SLAP1	Cell surface glycoprotein 1	Q06852	ENSCINP00000000374.3	
AIPGENE3003	SLAP1	Cell surface glycoprotein 1	Q06852	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE14454	SLN14	Schlafen family member 14	P0C7P3		
AIPGENE17849	SMD3	Small nuclear ribonucleoprotein Sm D3	P62323	ENSCINP00000024281.2	GO:0000245 GO:0000387 GO:0000932 GO:0000956 GO:0005688 GO:0006396 GO:0017070 GO:0033962 GO:0097526
AIPGENE3157	SNED1	Sushi, nidogen and EGF-like domain-containing protein 1	Q5ZQU0	ENSCINP00000030555.1	GO:0005509 GO:0007160
AIPGENE4545	SPAS1	Spermatogenesis-associated serine-rich protein 1	Q496A3	ENSCINP00000016015.3	
AIPGENE5872	SSPO	SCO-spondin	A2VEC9	ENSCINP00000035305.1	

AIPGENE13143	*STAB2*	*Stabilin-2*	*Q8WWQ8*	*ENSCINP00000032601.1*	*GO:0005509*
AIPGENE11122	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*A2AVA0*	*ENSCINP00000019994.3*	
AIPGENE13676	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*P0C6B8*	*ENSCINP00000025031.2*	
AIPGENE24130	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*Q4LDE5*	*ENSCINP00000007127.3*	*GO:0005509*
AIPGENE25129	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*P0C6B8*	*ENSCINP00000019994.3*	
AIPGENE5577	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*Q4LDE5*	*ENSCINP00000019994.3*	
AIPGENE9664	*SVEP1*	*Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1*	*P0C6B8*	*ENSCINP00000007127.3*	*GO:0005509*
AIPGENE26205	*SYRC*	*Arginine—tRNA ligase, cytoplasmic*	*P40329*	*ENSCINP00000015003.3*	*GO:0000166 GO:0004812 GO:0004814 GO:0005524 GO:0005737 GO:0006412 GO:0006418 GO:0006420 GO:0016874*
AIPGENE19101	*TAGL*	*Transgelin*	*P19966*	*ENSCINP00000025754.2*	
AIPGENE24599	*TBA1C*	*Tubulin alpha-1C chain*	*P68373*	*ENSCINP00000015767.3*	
AIPGENE24486	*TCPZ*	*T-complex protein 1 subunit zeta*	*Q5ZJ54*	*ENSCINP00000013348.3*	*GO:0000166 GO:0005524 GO:0005737 GO:0006457 GO:0051082*
AIPGENE11198	*TEPP*	*Testis, prostate and placenta-expressed protein*	*Q6URK8*	*ENSCINP00000030311.1*	
AIPGENE8447	*TF211*	*Transposon Tf2-11 polyprotein*	*Q9UR07*	*ENSCINP00000031157.1*	
AIPGENE21136	*TGFB2*	*Transforming growth factor beta-2*	*P30371*	*ENSCINP00000017469.3*	*GO:0005114 GO:0005125 GO:0005576 GO:0005615 GO:0007179 GO:0008083 GO:0010862 GO:0040007 GO:0042127 GO:0042981 GO:0043408 GO:0048468 GO:0060395*
AIPGENE18030	*TGMH*	*Hemocyte protein-glutamine gamma-glutamyltransferase*	*Q05187*	*ENSCINP00000028847.2*	*GO:0003810 GO:0018149 GO:0046872*
AIPGENE15507	*TLL1*	*Tolloid-like protein 1*	*O43897*	*ENSCINP00000007799.3*	*GO:0004222 GO:0005509 GO:0006508 GO:0008233 GO:0008237*

					GO:0008270 GO:0016787 GO:0046872
AIPGENE9660	TLL2	Tolloid-like protein 2	Q9WVM6	ENSCINP00000007799.3	GO:0004222 GO:0005509 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787 GO:0046872
AIPGENE19148	TM11D	Transmembrane protease serine 11D	O60235	ENSCINP00000025317.2	
AIPGENE20827	TM256	Transmembrane protein 256 homolog	A4K526		
AIPGENE445	TMED4	Transmembrane emp24 domain-containing protein 4	Q8R1V4	ENSCINP00000016857.3	
AIPGENE15399	TMED7	Transmembrane emp24 domain-containing protein 7	Q9Y3B3	ENSCINP00000003864.3	GO:0005789 GO:0006810 GO:0016020 GO:0016021
AIPGENE3070	TMM62	Transmembrane protein 62	Q0P6H9	ENSCINP00000013110.3	GO:0016020 GO:0016021 GO:0016787
AIPGENE1274	TMM79	Transmembrane protein 79	Q9BSE2		
AIPGENE2211	TMM98	Transmembrane protein 98	Q6INX1	ENSCINP00000034941.1	
AIPGENE13889	TNF10	Tumor necrosis factor ligand superfamily member 10	P50591	ENSCINP00000036550.1	
AIPGENE20516	TOP1	DNA topoisomerase 1	P41512	ENSCINP00000006314.3	GO:0003677 GO:0003917 GO:0003918 GO:0005694 GO:0005730 GO:0006260 GO:0006265 GO:0006338 GO:0007059 GO:0031298
AIPGENE6481	TREX2	Three prime repair exonuclease 2	Q9R1A9	ENSCINP00000032802.1	GO:0003676
AIPGENE3595	TRFE	Serotransferrin	P12346	ENSCINP00000008235.3	GO:0005576 GO:0016020 GO:0016021
AIPGENE551	TSAL	L-threonine ammonia-lyase	Q74FW6	ENSCINP00000019051.3	GO:0003941 GO:0004794 GO:0005737 GO:0006520 GO:0006565 GO:0006567 GO:0030170
AIPGENE14391	UBCD2	Ubiquitin-conjugating enzyme E2-24 kDa	P52485	ENSCINP00000014090.2	GO:0000166 GO:0005524
AIPGENE7518	UD12	UDP-glucuronosyltransferase 1-2	P70691		
AIPGENE20406	USH2A	Usherin	O75445	ENSCINP00000009921.3	GO:0016020 GO:0016021
AIPGENE26620	USH2A	Usherin	Q2QI47	ENSCINP00000010645.3	GO:0005102 GO:0030155 GO:0030334 GO:0045995
AIPGENE20205	USOM5	Uncharacterized skeletal organic matrix protein 5	B8VIU6		
AIPGENE26588	USOM5	Uncharacterized skeletal organic matrix protein 5	B8VIU6	ENSCINP00000024219.2	

AIPGENE3012	*VIAAT*	*Vesicular inhibitory amino acid transporter*	*Q6DIV6*	*ENSCINP00000017689.3*	*GO:0003333 GO:0015171 GO:0016020 GO:0016021*
AIPGENE2084	*VMA5A*	*von Willebrand factor A domain-containing protein 5A*	*O00534*	*ENSCINP00000012041.3*	*GO:0003950 GO:0006471*
AIPGENE19830	*VNT1B*	*Homeobox protein vent1B*	*Q9YH71*	*ENSCINP00000023047.2*	*GO:0003677 GO:0005634 GO:0006355 GO:0043565*
AIPGENE10344	*VWA2*	*von Willebrand factor A domain-containing protein 2*	*Q6DCQ6*	*ENSCINP00000031657.1*	
AIPGENE5655	*VWA7*	*von Willebrand factor A domain-containing protein 7*	*Q9JHA8*	*ENSCINP00000015858.3*	*GO:0005509*
AIPGENE659	*VWA8*	*von Willebrand factor A domain-containing protein 8*	*A3KMH1*	*ENSCINP00000018067.3*	*GO:0005524 GO:0016887*
AIPGENE13608	*WDR49*	*WD repeat-containing protein 49*	*Q8IV35*	*ENSCINP00000011378.2*	*GO:0005509 GO:0016020 GO:0016021*
AIPGENE15889	*WHRN*	*Whirlin*	*Q9P202*	*ENSCINP00000031643.1*	
AIPGENE18854	*XLRS1*	*Retinoschisin*	*Q9W6R5*	*ENSCINP00000032368.1*	*GO:0070492*
AIPGENE26095	*Y3136*	*Clavaminate synthase-like protein At3g21360*	*Q9LIG0*		
AIPGENE6960	*YCP9*	*Uncharacterized oxidoreductase C663.09c*	*Q7Z9I2*	*ENSCINP00000032251.1*	
AIPGENE22163	*YG31B*	*Transposon Ty3-G Gag-Pol polyprotein*	*Q99315*	*ENSCINP00000036259.1*	*GO:0003676 GO:0008270 GO:0015074*
AIPGENE2954	*YG31B*	*Transposon Ty3-G Gag-Pol polyprotein*	*Q99315*	*ENSCINP00000032038.1*	
AIPGENE26025	*YLL3*	*Smr domain-containing protein C11H11.03c*	*Q9UTP4*		
AIPGENE13903	*ZPP*	*ZP domain-containing protein*	*G8HTB6*	*ENSCINP00000013501.3*	*GO:0016020 GO:0016021*
AIPGENE7095	*ZPP*	*ZP domain-containing protein*	*G8HTB6*	*ENSCINP00000033680.1*	*GO:0016020 GO:0016021*
AIPGENE7102	*ZPP*	*ZP domain-containing protein*	*G8HTB6*	*ENSCINP00000033680.1*	*GO:0016020 GO:0016021*
AIPGENE843	*ZPP*	*ZP domain-containing protein*	*G8HTB6*	*ENSCINP00000033680.1*	*GO:0016020 GO:0016021*
AIPGENE10052	*Predicted Protein*	*Predicted Protein*			
AIPGENE10088	*Predicted Protein*	*Predicted Protein*			
AIPGENE10653	*Predicted protein*	*Predicted protein*	*A7SZR2*		
AIPGENE10676	*Predicted Protein*	*Predicted Protein*			
AIPGENE10818	*Predicted protein*	*Predicted protein*	*W4ZL70*		
AIPGENE10946	*Predicted protein*	*Predicted protein*	*K1QAD8*		
AIPGENE11064	*Predicted Protein*	*Predicted Protein*			
AIPGENE11080	*Predicted protein*	*Predicted protein*	*K1QFV0*	*ENSCINP00000031905.1*	
AIPGENE11332	*Predicted Protein*	*Predicted Protein*			
AIPGENE11392	*Predicted protein*	*Predicted protein*	*A7SFL7*		
AIPGENE11651	*Predicted protein*	*Predicted protein*	*A7SLY5*		
AIPGENE11857	*Predicted Protein*	*Predicted Protein*			
AIPGENE12058	*Predicted protein*	*Predicted protein*	*A0A087TAW6*	*ENSCINP00000024248.2*	*GO:0005432 GO:0006816 GO:0007154 GO:0016020 GO:0016021 GO:0035725 GO:0055085*
AIPGENE12150	*Predicted Protein*	*Predicted Protein*			
AIPGENE12226	*Predicted protein*	*Predicted protein*	*I1G3E1*		
AIPGENE12918	*Predicted protein*	*Predicted protein*	*A7SBQ7*		
AIPGENE13469	*Predicted protein*	*Predicted protein*	*A0A0D2WKZ7*		
AIPGENE1349	*Predicted protein*	*Predicted protein*	*A7RLA1*		
AIPGENE13694	*Predicted protein*	*Predicted protein*	*A7SH51*		

AIPGENE13984	*Predicted protein*	*Predicted protein*	*A0A0B7B5L7*		
AIPGENE14001	*Predicted protein*	*Predicted protein*	*A7T249*		
AIPGENE14083	*Predicted Protein*	*Predicted Protein*			
AIPGENE14403	*Predicted protein*	*Predicted protein*	*A7SKN1*	*ENSCINP00000021776.2*	
AIPGENE1471	*Predicted protein*	*Predicted protein*	*A7SY04*		
AIPGENE1482	*Predicted Protein*	*Predicted Protein*			
AIPGENE15550	*Predicted protein*	*Predicted protein*	*C3Z268*		
AIPGENE15663	*Predicted Protein*	*Predicted Protein*			
AIPGENE15795	*Predicted protein*	*Predicted protein*	*A7RNY8*		
AIPGENE15954	*Predicted protein*	*Predicted protein*	*A7S4R8*		
AIPGENE15987	*Predicted protein*	*Predicted protein*	*A7SRM2*		
AIPGENE16199	*Predicted Protein*	*Predicted Protein*			
AIPGENE16288	*Predicted Protein*	*Predicted Protein*			
AIPGENE16404	*Predicted protein*	*Predicted protein*	*A7SVZ2*		
AIPGENE16868	*Predicted protein*	*Predicted protein*	*A7RFD4*	*ENSCINP00000025194.2*	
AIPGENE17146	*Predicted protein*	*Predicted protein*	*A7S4R8*		
AIPGENE17179	*Predicted Protein*	*Predicted Protein*			
AIPGENE17299	*Predicted Protein*	*Predicted Protein*			
AIPGENE17543	*Predicted Protein*	*Predicted Protein*	*KXJ27203.1*	*GO:0005509*	
AIPGENE1755	*Predicted Protein*	*Predicted Protein*			
AIPGENE17878	*Predicted Protein*	*Predicted Protein*			
AIPGENE18192	*Predicted protein*	*Predicted protein*	*R7UVL8*		
AIPGENE18429	*Predicted Protein*	*Predicted Protein*			
AIPGENE18615	*Predicted Protein*	*Predicted Protein*			
AIPGENE19123	*Predicted protein*	*Predicted protein*	*A7SRW9*		
AIPGENE19480	*Predicted protein*	*Predicted protein*	*A7SPH6*	*ENSCINP00000032925.1*	
AIPGENE19594	*Predicted protein*	*Predicted protein*	*A7S0M0*		
AIPGENE19601	*Predicted protein*	*Predicted protein*	*V4BBZ1*		
AIPGENE19621	*Predicted protein*	*Predicted protein*	*A7SGB4*		
AIPGENE19727	*Predicted protein*	*Predicted protein*	*A7RJW3*		
AIPGENE20022	*Predicted protein*	*Predicted protein*	*M1SUN2*		
AIPGENE20142	*Predicted protein*	*Predicted protein*	*W4YRC9*		
AIPGENE20496	*Predicted Protein*	*Predicted Protein*			
AIPGENE20612	*Predicted protein*	*Predicted protein*	*C3Y7T5*	*ENSCINP00000003777.3*	
AIPGENE20703	*Predicted Protein*	*Predicted Protein*			
AIPGENE20797	*Predicted protein*	*Predicted protein*	*A7RYD2*		
AIPGENE20899	*Predicted protein*	*Predicted protein*	*A7SUN7*		
AIPGENE20957	*Predicted protein*	*Predicted protein*	*A7SUE7*		
AIPGENE21062	*Predicted protein*	*Predicted protein*	*A7SHP6*		
AIPGENE21286	*Predicted Protein*	*Predicted Protein*			
AIPGENE21863	*Predicted protein*	*Predicted protein*	*W4Y9V5*		
AIPGENE22000	*Predicted Protein*	*Predicted Protein*			
AIPGENE22430	*Predicted protein*	*Predicted protein*	*A7RTE7*		
AIPGENE23391	*Predicted protein*	*Predicted protein*	*A7S5Y5*	*ENSCINP00000026937.2*	*GO:0005634*
AIPGENE23393	*Predicted Protein*	*Predicted Protein*			
AIPGENE23394	*Predicted protein*	*Predicted protein*	*A7SJ52*		
AIPGENE23589	*Predicted protein*	*Predicted protein*	*A7SGY7*		
AIPGENE23764	*Predicted Protein*	*Predicted Protein*			
AIPGENE23808	*Predicted protein*	*Predicted protein*	*A7RTG8*		
AIPGENE2413	*Predicted protein*	*Predicted protein*	*A7SLX7*		
AIPGENE24756	*Predicted Protein*	*Predicted Protein*			
AIPGENE24771	*Predicted Protein*	*Predicted Protein*			
AIPGENE24791	*Predicted Protein*	*Predicted Protein*			
AIPGENE24795	*Predicted Protein*	*Predicted Protein*			
AIPGENE24832	*Predicted protein*	*Predicted protein*	*V4AH91*		
AIPGENE25047	*Predicted Protein*	*Predicted Protein*			
AIPGENE25048	*Predicted protein*	*Predicted protein*	*C3Z282*		
AIPGENE25162	*Predicted Protein*	*Predicted Protein*			
AIPGENE25215	*Predicted Protein*	*Predicted Protein*			
AIPGENE25239	*Predicted Protein*	*Predicted Protein*			
AIPGENE25284	*Predicted protein*	*Predicted protein*	*I1EVC9*		
AIPGENE26509	*Predicted protein*	*Predicted protein*	*K1R3J8*		
AIPGENE26956	*Predicted protein*	*Predicted protein*	*W4ZF57*		
AIPGENE27185	*Predicted protein*	*Predicted protein*	*A7RTP2*	*ENSCINP00000021776.2*	
AIPGENE2732	*Predicted Protein*	*Predicted Protein*			
AIPGENE27361	*Predicted Protein*	*Predicted Protein*			
AIPGENE27368	*Predicted protein*	*Predicted protein*	*A7S4E4*		
AIPGENE28010	*Predicted protein*	*Predicted protein*	*A7S9X2*		
AIPGENE28418	*Predicted protein*	*Predicted protein*	*A7RUV0*		
AIPGENE28604	*Predicted Protein*	*Predicted Protein*			
AIPGENE28810	*Predicted Protein*	*Predicted Protein*	*KXJ15997.1*	*GO:0003677 GO:0006352*	
AIPGENE28932	*Predicted Protein*	*Predicted Protein*			
AIPGENE28943	*Predicted protein*	*Predicted protein*	*A7SEN9*		
AIPGENE29057	*Predicted protein*	*Predicted protein*	*A7SSX4*		

AIPGENE29082	*Predicted Protein*	*Predicted Protein*		
AIPGENE29179	*Predicted Protein*	*Predicted Protein*		
AIPGENE3021	*Predicted protein*	*Predicted protein*	*A7S0X8*	
AIPGENE3111	*Predicted Protein*	*Predicted Protein*		
AIPGENE321	*Predicted protein*	*Predicted protein*	*A7RGE1*	
AIPGENE4202	*Predicted Protein*	*Predicted Protein*		
AIPGENE4274	*Predicted Protein*	*Predicted Protein*		
AIPGENE4311	*Predicted Protein*	*Predicted Protein*		
AIPGENE4475	*Predicted Protein*	*Predicted Protein*		
AIPGENE4911	*Predicted Protein*	*Predicted Protein*		
AIPGENE4937	*Predicted protein*	*Predicted protein*	*A7SFQ6*	
AIPGENE5651	*Predicted protein*	*Predicted protein*	*A7TCQ4*	
AIPGENE5667	*Predicted Protein*	*Predicted Protein*		
AIPGENE596	*Predicted Protein*	*Predicted Protein*		
AIPGENE5998	*Predicted Protein*	*Predicted Protein*		
AIPGENE6366	*Predicted protein*	*Predicted protein*	*I1FTA0*	*ENSCINP00000032925.1*
AIPGENE6504	*Predicted protein*	*Predicted protein*	*A7T1D0*	
AIPGENE6535	*Predicted protein*	*Predicted protein*	*V3ZKL5*	
AIPGENE6803	*Predicted Protein*	*Predicted Protein*		
AIPGENE7470	*Predicted Protein*	*Predicted Protein*		
AIPGENE7876	*Predicted protein*	*Predicted protein*	*F7ALQ1*	*ENSCINP00000010994.3*
AIPGENE8035	*Predicted protein*	*Predicted protein*	*A7RTG8*	
AIPGENE8046	*Predicted Protein*	*Predicted Protein*		
AIPGENE8081	*Predicted Protein*	*Predicted Protein*		
AIPGENE8504	*Predicted protein*	*Predicted protein*	*A7RFV5*	
AIPGENE8505	*Predicted Protein*	*Predicted Protein*		
AIPGENE9011	*Predicted protein*	*Predicted protein*	*A7SG11*	
AIPGENE9044	*Predicted Protein*	*Predicted Protein*		
AIPGENE9056	*Predicted protein*	*Predicted protein*	*A7SFP2*	
AIPGENE9074	*Predicted protein*	*Predicted protein*	*Q5I4C4*	
AIPGENE9090	*Predicted Protein*	*Predicted Protein*		
AIPGENE9134	*Predicted Protein*	*Predicted Protein*		
AIPGENE9225	*Predicted Protein*	*Predicted Protein*		
AIPGENE9239	*Predicted protein*	*Predicted protein*	*V4CSL7*	*ENSCINP00000034779.1*
AIPGENE9471	*Predicted protein*	*Predicted protein*	*A7RZR8*	
AIPGENE9726	*Predicted Protein*	*Predicted Protein*		
AIPGENE9753	*Predicted Protein*	*Predicted Protein*		

Table S1-4. Enriched GO terms for HIBA genes

E. diaphana geneID	Gene name	Discription	UniprotID	*C. intestinalis* geneID	GOID
AIPGENE22099	AASS	Alpha-aminoadipic semialdehyde synthase, mitochondrial	A8E657	ENSCINP00000019576.3	GO:0016491 GO:0055114
AIPGENE11605	ACAC	Acetyl-CoA carboxylase	P11029	ENSCINP00000005865.3	GO:0000166 GO:0003824 GO:0003989 GO:0004075 GO:0005524 GO:0006633 GO:0016874 GO:0046872
AIPGENE17854	ACEA	Isocitrate lyase	Q9K9H0		
AIPGENE20290	ACOD	Acyl-CoA desaturase	O62849	ENSCINP00000033408.1	
AIPGENE2600	ACTP2	Equinatoxin-2	P61914		
AIPGENE2649	ACTPC	Tenebrosin-C	P61915		
AIPGENE25748	AGRD1	Adhesion G-protein coupled receptor D1	A6QLU6	ENSCINP00000011544.3	GO:0004871 GO:0004888 GO:0004930 GO:0005509 GO:0005886 GO:0007155 GO:0007156 GO:0007165 GO:0007166 GO:0007186 GO:0016020 GO:0016021
AIPGENE5636	AL1L1	Cytosolic 10-formyltetrahydrofolate dehydrogenase	Q6GNL7	ENSCINP00000016529.3	GO:0003824 GO:0005737 GO:0006730 GO:0008152 GO:0009058 GO:0009258 GO:0016155 GO:0016491 GO:0016620 GO:0016742 GO:0019145 GO:0047105 GO:0055114
AIPGENE19640	AL3B1	Aldehyde dehydrogenase family 3 member B1	P43353	ENSCINP00000001020.3	GO:0004028 GO:0004029 GO:0004030 GO:0006081 GO:0008152 GO:0016491 GO:0016620 GO:0055114
AIPGENE17420	AMT1	Putative ammonium transporter 1	P54145	ENSCINP00000012012.3	GO:0005887 GO:0006810 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488
AIPGENE28006	ANAG	Alpha-N-acetylglucosaminidase	P54802	ENSCINP00000024455.2	
AIPGENE23315	AOSL	Allene oxide synthase-lipoxygenase protein	O16025	ENSCINP00000033534.1	GO:0016491 GO:0016702 GO:0046872 GO:0051213 GO:0055114
AIPGENE18609	ARID2	AT-rich interactive domain-containing protein 2	Q68CP9	ENSCINP00000013045.3	GO:0001664 GO:0005737 GO:0005764 GO:0006629 GO:0006665 GO:0008047 GO:0043085

AIPGENE6329	ARSB	Arylsulfatase B	P33727	ENSCINP00000007178.3	GO:0003824 GO:0008152 GO:0008484
AIPGENE18100	ATS18	A disintegrin and metalloproteinase with thrombospondin motifs 18	Q8TE60	ENSCINP00000018188.3	GO:0004222 GO:0005578 GO:0006508 GO:0008237 GO:0008270 GO:0031012
AIPGENE3533	ATTY	Tyrosine aminotransferase	Q8QZR1	ENSCINP00000000577.3	GO:0003824 GO:0004838 GO:0006520 GO:0008483 GO:0009058 GO:0009072 GO:0030170
AIPGENE1323	BCAS3	Breast carcinoma-amplified sequence 3 homolog	Q8CCN5	ENSCINP00000024644.2	
AIPGENE9338	BGLR	Beta-glucuronidase	O97524	ENSCINP00000032126.1	GO:0004553 GO:0004566 GO:0005764 GO:0005975 GO:0008152 GO:0016020 GO:0016021 GO:0016787 GO:0016798
AIPGENE9365	BGLR	Beta-glucuronidase	Q4FAT7	ENSCINP00000010614.3	GO:0004553 GO:0004566 GO:0005764 GO:0005975 GO:0008152 GO:0016787 GO:0016798 GO:0043231
AIPGENE1836	BPI	Bactericidal permeability-increasing protein	P17213	ENSCINP00000033806.1	
AIPGENE2901	CAH2	Carbonic anhydrase 2	Q8UWA5	ENSCINP00000003284.3	GO:0004089 GO:0006730 GO:0008270 GO:0046872
AIPGENE16849	CATB	Cathepsin B	A1E295	ENSCINP00000031460.1	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008234 GO:0050790 GO:0051603
AIPGENE26156	CATL	Cathepsin L	Q26636	ENSCINP00000015731.3	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008233 GO:0008234 GO:0016787 GO:0051603
AIPGENE1298	CD63	CD63 antigen	Q76B49	ENSCINP00000031139.1	
AIPGENE1271	CFAD	Counting factor associated protein D	Q54TR1	ENSCINP00000015736.3	GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0008234 GO:0051603
AIPGENE11370	CGL	Cystathionine gamma-lyase	Q55DV9	ENSCINP00000025947.2	GO:0003824 GO:0030170
AIPGENE7524	CGT	2-hydroxyacylsphingosine 1-beta-galactosyltransferase	Q16880		
AIPGENE3160	CHST1	Carbohydrate sulfotransferase 1	Q6DBY9	ENSCINP00000022864.2	GO:0001517 GO:0006790 GO:0008146

					GO:0016021 GO:0016740
AIPGENE7759	CNG	Cyclic nucleotide-gated cation channel	P55934	ENSCINP00000014322.3	GO:0005216 GO:0005249 GO:0005887 GO:0006810 GO:0006811 GO:0006813 GO:0016020 GO:0016021 GO:0034220 GO:0042391 GO:0055085 GO:0071805
AIPGENE4172	CNN1	Calponin-1	P26932	ENSCINP00000024676.2	GO:0003779 GO:0005516 GO:0031032
AIPGENE14130	COLA1	Collagen alpha-1(XXI) chain	Q96P44	ENSCINP00000016093.3	GO:0005201 GO:0005578 GO:0005581
AIPGENE19050	COLA1	Collagen alpha-1(XXI) chain	Q96P44	ENSCINP00000016017.3	
AIPGENE1554	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	P30437	ENSCINP00000032679.1	
AIPGENE23070	CP17A	Steroid 17-alpha-hydroxylase/17,20 lyase	O73853	ENSCINP00000022528.2	
AIPGENE28157	CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial	P31327	ENSCINP00000018277.3	GO:0000050 GO:0003824 GO:0004070 GO:0004087 GO:0004088 GO:0004151 GO:0005524 GO:0005737 GO:0006207 GO:0006526 GO:0006807 GO:0046872
AIPGENE21331	CPT1A	Carnitine O-palmitoyltransferase 1, liver isoform	P50416	ENSCINP00000007072.3	GO:0016020 GO:0016021 GO:0016746
AIPGENE13178	CTP5A	Contactin-associated protein like 5-1	Q0V8T9	ENSCINP00000010540.3	
AIPGENE5888	CTRB	Chymotrypsinogen B	P00767	ENSCINP00000031261.1	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE17825	CTX1	Toxin CrTX-A	Q9GV72		
AIPGENE17834	CTX1	Toxin CrTX-A	Q9GV72		
AIPGENE5556	CYS2	Probable serine-O-acetyltransferase cys2	Q10341	ENSCINP00000033915.1	GO:0009001 GO:0016787 GO:0019344
AIPGENE22204	CYSP	Cathepsin B-like cysteine proteinase	P25792	ENSCINP00000019058.3	GO:0004177 GO:0004197 GO:0005615 GO:0005764 GO:0006508 GO:0006919 GO:0006955 GO:0008233 GO:0008234 GO:0008656 GO:0010952 GO:0016787 GO:0051603 GO:2001235
AIPGENE6491	DHCR7	7-dehydrocholesterol reductase	Q7SXF1	ENSCINP00000014788.3	GO:0006695 GO:0016020 GO:0016021 GO:0016628 GO:0030176 GO:0047598 GO:0055114

AIPGENE24026	DHE3	Glutamate dehydrogenase	P94598	ENSCINP00000019611.3	GO:0006520 GO:0016491 GO:0016639 GO:0055114
AIPGENE27618	DHE3	Glutamate dehydrogenase	P94598	ENSCINP00000019611.3	GO:0006520 GO:0016491 GO:0016639 GO:0055114
AIPGENE20433	DIRA2	GTP-binding protein Di-Ras2	Q5R6S2	ENSCINP00000015095.3	GO:0000166 GO:0005525 GO:0005622 GO:0007165 GO:0007264 GO:0016020
AIPGENE20966	DIRC2	Disrupted in renal carcinoma protein 2 homolog	Q6GNV7	ENSCINP00000023309.2	
AIPGENE15640	DMBT1	Deleted in malignant brain tumors 1 protein	Q9UGM3	ENSCINP00000015917.3	GO:0004252 GO:0006508 GO:0008233 GO:0008236 GO:0016787
AIPGENE3208	DPF1	Zinc finger protein neuro-d4	Q9QX66	ENSCINP00000033757.1	GO:0008270 GO:0046872
AIPGENE3939	DSCAM	Down syndrome cell adhesion molecule	O60469	ENSCINP00000015860.3	GO:0005509
AIPGENE24155	DYH3	Dynein heavy chain 3, axonemal	Q8BW94	ENSCINP00000031667.1	GO:0003777 GO:0007018
AIPGENE18149	ERMP1	Endoplasmic reticulum metallopeptidase 1	Q3UVK0	ENSCINP00000035901.1	GO:0016020 GO:0016021
AIPGENE19098	FA46C	Protein FAM46C	Q7ZUP1	ENSCINP00000026866.2	
AIPGENE14819	FAD5	Delta(5) fatty acid desaturase	O74212	ENSCINP00000004431.3	GO:0016020 GO:0016021 GO:0020037 GO:0046872
AIPGENE17789	FAMT	Probable fatty acid methyltransferase Rv3720	O69687		
AIPGENE11718	FAXC	Failed axon connections homolog	F7E235	ENSCINP00000025230.2	GO:0004364 GO:0005737 GO:0006749
AIPGENE11705	FBN2	Fibrillin-2	P35556	ENSCINP00000010391.3	GO:0005201 GO:0005509 GO:0005578
AIPGENE27824	FBP1	Fibropellin-1	P10079	ENSCINP00000014588.3	
AIPGENE22541	FGFR2	Fibroblast growth factor receptor 2	Q91286	ENSCINP00000001598.3	GO:0004672 GO:0004713 GO:0005007 GO:0005524 GO:0006468 GO:0008284 GO:0008543 GO:0016020 GO:0016021 GO:0018108
AIPGENE12192	FMO2	Dimethylaniline monooxygenase [N-oxide-forming] 2	Q8HZ70	ENSCINP00000011417.3	GO:0004497 GO:0004499 GO:0005783 GO:0005789 GO:0016020 GO:0016021 GO:0016491 GO:0017144 GO:0031090 GO:0043231 GO:0050660 GO:0050661 GO:0055114
AIPGENE21768	FMO5	Dimethylaniline monooxygenase [N-oxide-forming] 5	Q8K4C0	ENSCINP00000011417.3	GO:0004497 GO:0004499 GO:0005783 GO:0005789 GO:0016020 GO:0016021

					GO:0016491 GO:0017144 GO:0031090 GO:0043231 GO:0050660 GO:0050661 GO:0055114
AIPGENE26282	FOXO	Forkhead box protein O	O16850	ENSCINP00000035316.1	GO:0003677 GO:0003700 GO:0005634 GO:0006351 GO:0006355 GO:0043565
AIPGENE6958	FP	Fibronectin type III domain-containing protein	B8VIW9		
AIPGENE22443	FRIS	Soma ferritin	P42577	ENSCINP00000035294.1	
AIPGENE14754	FRRS1	Putative ferric-chelate reductase 1	Q6INU7	ENSCINP00000017029.3	
AIPGENE9887	FSTL5	Follistatin-related protein 5	Q8N475	ENSCINP00000025719.2	
AIPGENE16012	FUCO_BRAFL	Alpha-L-fucosidase2	C3YWU0	ENSCINP00000015524.3	
AIPGENE10000	GALK2	N-acetylgalactosamine kinase	Q5XIG6	ENSCINP00000028204.2	GO:0000166 GO:0004335 GO:0005524 GO:0005737 GO:0006012 GO:0008152 GO:0016301 GO:0016310 GO:0016740 GO:0016773 GO:0046835
AIPGENE15776	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000001814.3	GO:0005576
AIPGENE25324	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9CYL5	ENSCINP00000001814.3	GO:0005576
AIPGENE8921	GAPR1	Golgi-associated plant pathogenesis-related protein 1	Q9H4G4	ENSCINP00000010145.3	GO:0005576
AIPGENE1512	GBA3	Cytosolic beta-glucosidase	Q5RF65	ENSCINP00000010342.3	GO:0004553 GO:0005975 GO:0008422 GO:1901657
AIPGENE1160	GBP3	Guanylate-binding protein 3	Q9H0R5	ENSCINP00000032288.1	
AIPGENE23694	GGLO	L-gulonolactone oxidase	P58710		
AIPGENE3097	GNPAT	Dihydroxyacetone phosphate acyltransferase	O15228	ENSCINP00000018520.3	GO:0008152 GO:0008374 GO:0016746 GO:0044255
AIPGENE182	GNS	N-acetylglucosamine-6-sulfatase	P50426	ENSCINP00000019163.3	
AIPGENE5657	GPX1	Glutathione peroxidase	Q00277	ENSCINP00000035821.1	GO:0004601 GO:0004602 GO:0006979 GO:0016491 GO:0055114 GO:0098869
AIPGENE1907	GRM8	Metabotropic glutamate receptor 8	P47743	ENSCINP00000025414.2	
AIPGENE20480	GST	Glutathione S-transferase	P46428	ENSCINP00000000153.3	
AIPGENE12082	GTR1	Solute carrier family 2, facilitated glucose transporter member 1	P20303	ENSCINP00000003472.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE2706	GTR8	Solute carrier family 2,	Q9NY64	ENSCINP00000002435.3	GO:0005215

		facilitated glucose transporter member 8			GO:0005351 GO:0005355 GO:0005887 GO:0006810 GO:0015992 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0035428 GO:0046323 GO:0055085 GO:1904659
AIPGENE6170	GTR8	Solute carrier family 2, facilitated glucose transporter member 8	Q9JJZ1	ENSCINP00000019539.3	GO:0005351 GO:0005355 GO:0005887 GO:0006810 GO:0015992 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0035428 GO:0046323 GO:0055085 GO:1904659
AIPGENE20289	HEX	Beta-hexosaminidase	Q04786	ENSCINP00000008969.3	GO:0004553 GO:0004563 GO:0005975
AIPGENE24739	HRH2	Histamine H2 receptor	P47747	ENSCINP00000009895.3	GO:0004930 GO:0007186 GO:0016020 GO:0016021
AIPGENE22432	HSP31	Glyoxalase 3	Q5AF03		
AIPGENE21738	HSP7C	Heat shock cognate 71 kDa protein	Q71U34	ENSCINP00000019652.3	
AIPGENE17792	HUTH	Histidine ammonia-lyase	P42357	ENSCINP00000013857.3	GO:0003824 GO:0004397 GO:0005737 GO:0006547 GO:0006548 GO:0016829 GO:0016841 GO:0019556 GO:0019557
AIPGENE10180	IFIH1	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	ENSCINP00000031621.1	
AIPGENE18569	IFIH1	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	ENSCINP00000031621.1	
AIPGENE25676	INO1A	Inositol-3-phosphate synthase 1-A	Q7ZXY0		
AIPGENE20336	INSI2	Insulin-induced gene 2 protein	Q91WG1		
AIPGENE24707	IRF2	Interferon regulatory factor 2	P14316	ENSCINP00000005767.3	GO:0000975 GO:0003700 GO:0006355
AIPGENE27863	KARG	Arginine kinase	O15990	ENSCINP00000017993.3	
AIPGENE7287	KLC	Kinesin light chain	P46824	ENSCINP00000034474.1	
AIPGENE9819	KLC	Kinesin light chain	P46822	ENSCINP00000034474.1	
AIPGENE26876	KLHL3	Kelch-like protein 3	E0CZ16	ENSCINP00000008089.3	GO:0004842 GO:0031463 GO:0042787
AIPGENE13828	LACS8	Long chain acyl-CoA synthetase 8	Q9SJD4	ENSCINP00000005887.3	GO:0003824 GO:0008152
AIPGENE23106	LGMN	Legumain	Q95M12	ENSCINP00000006801.3	GO:0004197 GO:0005773 GO:0006508 GO:0006624 GO:0008233 GO:0051603
AIPGENE26593	LIP1	Lipase ZK262.3	Q9XTR8		
AIPGENE8872	LIPR2	Pancreatic lipase-	P54318	ENSCINP00000014633.3	GO:0004806

		related protein 2			GO:0005576 GO:0006629 GO:0052689
AIPGENE13786	LSD2	Lipid storage droplets surface-binding protein 2	Q9VXY7		
AIPGENE13692	LYAG	Lysosomal alpha-glucosidase	P10253	ENSCINP00000008112.3	GO:0003824 GO:0004553 GO:0005975 GO:0008152 GO:0016787 GO:0016798 GO:0030246
AIPGENE6132	MA2B1	Lysosomal alpha-mannosidase	Q60HE9	ENSCINP00000003992.3	GO:0003824 GO:0004553 GO:0004559 GO:0005975 GO:0006013 GO:0006517 GO:0008152 GO:0008270 GO:0015923 GO:0016787 GO:0016798 GO:0030246 GO:0046872
AIPGENE6148	MA2B1	Lysosomal alpha-mannosidase	Q60HE9	ENSCINP00000003992.3	GO:0003824 GO:0004553 GO:0004559 GO:0005975 GO:0006013 GO:0006517 GO:0008152 GO:0008270 GO:0015923 GO:0016787 GO:0016798 GO:0030246 GO:0046872
AIPGENE4828	MCA7	Metacaspase-7	Q6XPT5		
AIPGENE15960	MCAT	Mitochondrial carnitine/acylcarnitine carrier protein	Q9Z2Z6	ENSCINP00000016773.3	GO:0003735 GO:0006412 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE9553	MET7B	Methyltransferase-like protein 7B	Q6UX53	ENSCINP00000027827.2	GO:0005737 GO:0005829 GO:0008757 GO:0030791 GO:0032259
AIPGENE2103	MFGM	Lactadherin	P70490	ENSCINP00000006532.3	GO:0005509
AIPGENE20554	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE20557	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE26741	MFS12	Major facilitator superfamily domain-containing protein 12	Q3U481	ENSCINP00000031725.1	GO:0016020 GO:0016021
AIPGENE14711	MFSD6	Major facilitator superfamily domain-containing protein 6	Q8CBH5	ENSCINP00000006447.3	GO:0016020 GO:0016021
AIPGENE5656	MGLL	Monoglyceride lipase	Q99685	ENSCINP00000020690.3	
AIPGENE14222	MIB1	E3 ubiquitin-protein ligase MIB1	Q86YT6	ENSCINP00000001354.3	GO:0004842 GO:0005737 GO:0006897 GO:0007219 GO:0008270 GO:0016567 GO:0046872
AIPGENE14234	MIB2	E3 ubiquitin-protein ligase MIB2	Q5ZIJ9	ENSCINP00000001354.3	GO:0004842 GO:0005737

					GO:0006897 GO:0007219 GO:0008270 GO:0016567 GO:0046872
AIPGENE15748	MIP	Lens fiber major intrinsic protein	Q6J8I9	ENSCINP00000011466.3	GO:0005215 GO:0005887 GO:0006810 GO:0006833 GO:0009992 GO:0015250 GO:0015254 GO:0015793 GO:0016020 GO:0016021 GO:0034220
AIPGENE19158	MOG2A	2-acylglycerol O-acyltransferase 2-A	Q2KHS5	ENSCINP00000004063.3	GO:0016747
AIPGENE1928	MOT10	Monocarboxylate transporter 10	Q3U9N9	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE1268	MOT5	Monocarboxylate transporter 5	Q8R0M8	ENSCINP00000026236.2	
AIPGENE3158	MOT8	Monocarboxylate transporter 8	Q8K1P8	ENSCINP00000007886.2	GO:0005215 GO:0016020 GO:0016021 GO:0034220 GO:0055085
AIPGENE15605	MPEG1	Macrophage-expressed gene 1 protein	Q9WV57		
AIPGENE1799	MRC1	Macrophage mannose receptor 1	Q61830	ENSCINP00000004176.3	GO:0016020 GO:0016021
AIPGENE4554	MUC2	Mucin-2 (Fragment)	Q62635	ENSCINP00000035305.1	
AIPGENE8101	MYOF	Myoferlin	B3DLH6	ENSCINP00000013715.3	GO:0016020 GO:0016021
AIPGENE3953	N2F1A	Nuclear receptor subfamily 2 group F member 1-A	Q06725	ENSCINP00000004102.3	GO:0003677 GO:0003700 GO:0003707 GO:0004879 GO:0005634 GO:0006351 GO:0006355 GO:0008270 GO:0030522 GO:0043401 GO:0043565 GO:0046872
AIPGENE13027	NAGAB	Alpha-N-acetylgalactosaminidase	Q90744	ENSCINP00000031972.1	GO:0003824 GO:0004553 GO:0004557 GO:0005737 GO:0005975 GO:0008152 GO:0016139 GO:0016787 GO:0016798
AIPGENE13031	NAGAB	Alpha-N-acetylgalactosaminidase	Q90744	ENSCINP00000031972.1	GO:0003824 GO:0004553 GO:0004557 GO:0005737 GO:0005975 GO:0008152 GO:0016139 GO:0016787 GO:0016798
AIPGENE17549	NAS4	Zinc metalloproteinase nas-4	P55112	ENSCINP00000030500.1	GO:0004222 GO:0006508 GO:0008233 GO:0008237 GO:0008270 GO:0016787 GO:0046872

AIPGENE8386	NAT16	Probable N-acetyltransferase 16	Q8N8M0		
AIPGENE25378	NCA11	Neural cell adhesion molecule 1-A	P16170	ENSCINP00000015860.3	GO:0005509
AIPGENE24468	NCL1	B-box type zinc finger protein ncl-1	P34611		
AIPGENE5315	NEP	Neprilysin	P08049	ENSCINP00000011815.3	GO:0004222 GO:0006508 GO:0008237
AIPGENE25764	NETR	Neurotrypsin	O08762	ENSCINP00000013067.3	GO:0005044 GO:0005507 GO:0006898 GO:0016020 GO:0016641 GO:0055114
AIPGENE18582	NFH	Neurofilament heavy polypeptide	P19246		
AIPGENE1175	NGB	Neuroglobin	Q99JA8	ENSCINP00000015555.3	GO:0005344 GO:0006810 GO:0015671 GO:0019825 GO:0020037
AIPGENE20032	NPC2	Epididymal secretory protein E1	P79345	ENSCINP00000030926.1	
AIPGENE22473	NPC2	Epididymal secretory protein E1	P79345	ENSCINP00000030926.1	
AIPGENE22527	NPC2	Protein NPC2 homolog	Q9VQ62	ENSCINP00000021760.2	
AIPGENE3074	NPC2	Protein NPC2 homolog	Q9VQ62	ENSCINP00000021760.2	
AIPGENE26780	NPHP3	Nephrocystin-3	Q7Z494	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE4737	NPHP3	Nephrocystin-3	Q7Z494	ENSCINP00000034474.1	
AIPGENE16407	NRF6	Nose resistant to fluoxetine protein 6	Q09225	ENSCINP00000013754.3	GO:0016020 GO:0016021 GO:0016747
AIPGENE11352	NWD1	NACHT domain- and WD repeat-containing protein 1	A6H603	ENSCINP00000035575.1	
AIPGENE17506	OCTL	Organic cation transporter-like protein	Q95R48	ENSCINP00000001843.3	GO:0005215 GO:0016020 GO:0016021 GO:0022857 GO:0022891 GO:0055085
AIPGENE6868	OIT3	Oncoprotein-induced transcript 3 protein	Q29RU2		
AIPGENE4696	PA2	Phospholipase A2	D2X8K2	ENSCINP00000028223.2	
AIPGENE24257	PAN2	PAB-dependent poly(A)-specific ribonuclease subunit PAN2	Q504Q3	ENSCINP00000000762.3	GO:0000289 GO:0000932 GO:0003676 GO:0004535 GO:0031251 GO:0090503
AIPGENE16667	PAPL	Iron/zinc purple acid phosphatase-like protein	A5D6U8		
AIPGENE3484	PCP	Lysosomal Pro-X carboxypeptidase	Q5RBU7	ENSCINP00000014308.3	GO:0003085 GO:0004185 GO:0006508 GO:0008236 GO:0008239 GO:0043535 GO:0060055
AIPGENE11495	PHR	Deoxyribodipyrimidine photo-lyase	Q28811		
AIPGENE11501	PISD	Phosphatidylserine decarboxylase proenzyme	Q5R8I8	ENSCINP00000036014.1	GO:0004609 GO:0008654
AIPGENE17365	PKHL1	Fibrocystin-L	Q80ZA4	ENSCINP00000032083.1	
AIPGENE762	PLBL2	Putative phospholipase B-like 2	Q2KIY5	ENSCINP00000021735.2	
AIPGENE11529	PLD2	Phospholipase D2	O14939	ENSCINP00000026201.2	
AIPGENE7602	PPT1	Palmitoyl-protein thioesterase 1	Q8HXW6	ENSCINP00000016630.3	GO:0002084 GO:0005764

					GO:0006898 GO:0008474 GO:0098599
AIPGENE8936	PRPX	Pathogen-related protein	P16273		
AIPGENE27212	PSA	Puromycin-sensitive aminopeptidase	P55786	ENSCINP00000014130.3	GO:0005737 GO:0005886 GO:0006508 GO:0008237 GO:0008270 GO:0042277 GO:0043171 GO:0070006
AIPGENE9002	PXMP2	Peroxisomal membrane protein 2	Q07066	ENSCINP00000026931.2	
AIPGENE20314	RAB32	Ras-related protein Rab-32	Q9CZE3	ENSCINP00000013519.3	GO:0005525 GO:0005739 GO:0005802 GO:0007264 GO:0032438 GO:0042470
AIPGENE16826	RDHE2	Epidermal retinol dehydrogenase 2	Q7TQA3	ENSCINP00000002227.3	GO:0016020 GO:0016021
AIPGENE25745	REG1A	Lithostathine-1-alpha	P05451	ENSCINP00000030082.1	
AIPGENE2133	RET	Proto-oncogene tyrosine-protein kinase receptor Ret	P35546	ENSCINP00000007167.3	GO:0004672 GO:0004713 GO:0004714 GO:0005524 GO:0006468 GO:0007169 GO:0016020 GO:0016021 GO:0018108
AIPGENE18105	RHBG	Ammonium transporter Rh type B	Q7T070	ENSCINP00000018176.3	GO:0005887 GO:0008519 GO:0015695 GO:0015696 GO:0016020 GO:0016021 GO:0019740 GO:0072488
AIPGENE3139	RNOY	Ribonuclease Oy	Q7M456	ENSCINP00000030837.1	
AIPGENE18289	RTJK	RNA-directed DNA polymerase from mobile element jockey	P21328	ENSCINP00000031852.1	
AIPGENE1279	SLC15A4	Solute carrier family 15 member 4	Q68F72	ENSCINP00000014849.3	GO:0005215 GO:0006810 GO:0006857 GO:0016020 GO:0016021
AIPGENE27783	SLC15A4	Solute carrier family 15 member 4	Q68F72	ENSCINP00000000708.3	GO:0005215 GO:0006810 GO:0006857 GO:0016020 GO:0016021
AIPGENE27827	SLC15A4	Solute carrier family 15 member 4	O09014	ENSCINP00000033357.1	
AIPGENE20198	S1C2A	Sulfotransferase 1C2A	Q9WUW9	ENSCINP00000035987.1	
AIPGENE26569	S238B	Solute carrier family 25 member 38-B	P0CAT2	ENSCINP00000018439.3	
AIPGENE24231	S23A2	Solute carrier family 23 member 2	B0JZG0	ENSCINP00000007695.3	GO:0005215 GO:0006810 GO:0016020 GO:0016021 GO:0055085
AIPGENE8501	S26A4	Pendrin	Q9R154	ENSCINP00000022644.2	
AIPGENE3906	S6A11	Sodium- and chloride-dependent GABA transporter 3	P48066	ENSCINP00000017435.3	
AIPGENE6454	S6A13	Sodium- and chloride-dependent GABA transporter 2	P31646	ENSCINP00000017435.3	
AIPGENE17806	SC5A3	Sodium/myo-inositol cotransporter	P31637	ENSCINP00000014618.3	

AIPGENE23673	SC5D	Lathosterol oxidase	O88822	ENSCINP00000027428.2	
AIPGENE3555	SCPDL	Saccharopine dehydrogenase-like oxidoreductase	Q8NBX0	ENSCINP00000023327.2	GO:0016020 GO:0016021 GO:0016491 GO:0055114
AIPGENE21060	SCRK	Putative fructokinase	O05510		
AIPGENE17639	SEM5B	Semaphorin-5B	Q9P283	ENSCINP00000015858.3	GO:0005509
AIPGENE19729	SLD1	Delta(8)-fatty-acid desaturase	Q43469	ENSCINP00000019998.3	
AIPGENE2530	SPDEF	SAM pointed domain-containing Ets transcription factor	Q9WTP3	ENSCINP00000024831.2	GO:0000981 GO:0003677 GO:0003700 GO:0005634 GO:0006355 GO:0006357 GO:0030154 GO:0043565
AIPGENE11862	STAR9	StAR-related lipid transfer protein 9	Q9P2P6	ENSCINP00000000065.3	GO:0008289
AIPGENE13021	SYT7	Synaptotagmin-7	O43581	ENSCINP00000001641.3	GO:0005509 GO:0005544 GO:0005886 GO:0006887 GO:0006906 GO:0016020 GO:0017158 GO:0019905 GO:0030276 GO:0048791 GO:0098793
AIPGENE12941	SYT9	Synaptotagmin-9	Q86SS6	ENSCINP00000001641.3	GO:0005509 GO:0005544 GO:0005886 GO:0006887 GO:0006906 GO:0016020 GO:0017158 GO:0019905 GO:0030276 GO:0048791 GO:0098793
AIPGENE3467	T53I2	Tumor protein p53-inducible nuclear protein 2	Q8IXH6		
AIPGENE13971	TC1A	Transposable element Tc1 transposase	P03934		
AIPGENE18026	TGM1	Protein-glutamine gamma-glutamyltransferase K	P23606	ENSCINP00000010148.3	GO:0003810 GO:0018149 GO:0046872
AIPGENE11194	TM189	Transmembrane protein 189	Q99LQ7	ENSCINP00000018140.3	GO:0005737 GO:0016020 GO:0016021 GO:0016567 GO:0031625 GO:0061630
AIPGENE3242	TM53B	Transmembrane protein 53-B	Q6DJC8	ENSCINP00000004026.3	GO:0016020 GO:0016021
AIPGENE203	TPP1	Tripeptidyl-peptidase 1	Q9EQV6		
AIPGENE18406	TRET1	Facilitated trehalose transporter Tret1	A9ZSY3	ENSCINP00000014757.3	
AIPGENE18099	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000018424.3	
AIPGENE24471	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE24487	TTC28	Tetratricopeptide repeat protein 28	Q96AY4	ENSCINP00000035104.1	GO:0005813 GO:0007346 GO:0030496
AIPGENE27733	TYRO	Tyrosinase	P06845	ENSCINP00000008838.3	GO:0008152 GO:0016491 GO:0046872 GO:0055114
AIPGENE20117	UBC23	Probable ubiquitin-conjugating enzyme E2	Q9ZVX1	ENSCINP00000023645.2	

		23			
AIPGENE15989	USOM3	Uncharacterized skeletal organic matrix protein 3 (Fragment)	B8RJM0		
AIPGENE5267	USOM5	Uncharacterized skeletal organic matrix protein 5	B8VIU6		
AIPGENE20926	VGFR2	Vascular endothelial growth factor receptor 2	P35918	ENSCINP00000015860.3	GO:0005509
AIPGENE6809	VITRN	Vitrin	Q6UXI7	ENSCINP00000031657.1	
AIPGENE17551	VKT1	Kunitz-type protease inhibitor AXPI-I	P81547	ENSCINP00000033091.1	GO:0004867 GO:0010951
AIPGENE18593	VLDLR	Very low-density lipoprotein receptor	P98166	ENSCINP00000031268.1	GO:0005509 GO:0016020 GO:0016021
AIPGENE10923	Y928	Putative protein methyltransferase MJ0928	Q58338	ENSCINP00000025027.2	GO:0001510 GO:0008168 GO:0009452
AIPGENE4562	YH24	Putative aminopeptidase W07G4.4	Q27245	ENSCINP00000016449.3	GO:0004177 GO:0005622 GO:0005737 GO:0006508 GO:0008235 GO:0019538 GO:0030145
AIPGENE15713	YL126	Putative glutamine amidotransferase YLR126C	Q12288		
AIPGENE11737	YL8A	Uncharacterized methyltransferase-like C25B8.10	Q9UTA8	ENSCINP00000017019.3	
AIPGENE13023	ZCH24	Zinc finger CCHC domain-containing protein 24	B2RVL6	ENSCINP00000010380.3	GO:0003676 GO:0008270
AIPGENE22382	ZN665	Zinc finger protein 665	Q5R8X1	ENSCINP00000001925.3	GO:0003676 GO:0046872
AIPGENE17949	ZNT1	Zinc transporter 1	Q9Y6M5	ENSCINP00000022636.2	
AIPGENE10556	Predicted Protein	Predicted Protein			
AIPGENE10738	Predicted Protein	Predicted Protein			
AIPGENE10743	Predicted protein	Predicted protein	V4A2K2		
AIPGENE10749	Predicted protein	Predicted protein	A7SA07		
AIPGENE10853	Predicted Protein	Predicted Protein			
AIPGENE1092	Predicted Protein	Predicted Protein			
AIPGENE11316	Predicted protein	Predicted protein	A7SUN8		
AIPGENE11357	Predicted protein	Predicted protein	A7SUN8		
AIPGENE11363	Predicted protein	Predicted protein	D2VQ64		
AIPGENE12286	Predicted Protein	Predicted Protein			
AIPGENE12420	Predicted Protein	Predicted Protein			
AIPGENE12482	Predicted Protein	Predicted Protein			
AIPGENE12722	Predicted Protein	Predicted Protein			
AIPGENE13174	Predicted Protein	Predicted Protein			
AIPGENE13688	Predicted protein	Predicted protein	A7S3A5	ENSCINP00000031518.1	
AIPGENE13917	Predicted Protein	Predicted Protein			
AIPGENE14095	Predicted Protein	Predicted Protein			
AIPGENE14493	Predicted Protein	Predicted Protein			
AIPGENE15114	Predicted Protein	Predicted Protein			
AIPGENE16249	Predicted protein	Predicted protein	R7UA65		
AIPGENE16293	Predicted Protein	Predicted Protein			
AIPGENE16299	Predicted protein	Predicted protein	E7EXN5	ENSCINP00000034291.1	
AIPGENE16386	Predicted protein	Predicted protein	B3RVQ7	ENSCINP00000035970.1	GO:0016020 GO:0016021
AIPGENE16517	Predicted protein	Predicted protein	A7RQD5		
AIPGENE16932	Predicted Protein	Predicted Protein			
AIPGENE17203	Predicted Protein	Predicted Protein			
AIPGENE17481	Predicted Protein	Predicted Protein			
AIPGENE17491	Predicted protein	Predicted protein	T1J4M1		
AIPGENE1754	Predicted protein	Predicted protein	A7SG99		
AIPGENE17637	Predicted protein	Predicted protein	Q0ZST9		
AIPGENE17686	Predicted Protein	Predicted Protein			
AIPGENE18567	Predicted Protein	Predicted Protein			
AIPGENE18767	Predicted Protein	Predicted Protein			
AIPGENE19267	Predicted Protein	Predicted Protein			
AIPGENE19468	Predicted Protein	Predicted Protein			
AIPGENE19608	Predicted protein	Predicted protein	C3ZHH9		

AIPGENE2021	Predicted protein	Predicted protein	A7SD04		
AIPGENE20726	Predicted protein	Predicted protein	C3Z8R0	ENSCINP00000032921.1	GO:0005802 GO:0016020 GO:0016021
AIPGENE21014	Predicted protein	Predicted protein	A7SIL6		
AIPGENE21365	Predicted protein	Predicted protein	A7SA37	ENSCINP00000012169.3	GO:0004930 GO:0004935 GO:0005887 GO:0007186 GO:0007267 GO:0016020 GO:0016021 GO:0071880
AIPGENE21580	Predicted protein	Predicted protein	C3ZYA3		
AIPGENE22726	Predicted Protein	Predicted Protein			
AIPGENE23311	Predicted protein	Predicted protein	A7SGG2		
AIPGENE23942	Predicted Protein	Predicted Protein			
AIPGENE24338	Predicted protein	Predicted protein	A7RSY4		
AIPGENE25157	Predicted protein	Predicted protein	C3XUM9		
AIPGENE26147	Predicted Protein	Predicted Protein			
AIPGENE26273	Predicted Protein	Predicted Protein			
AIPGENE26386	Predicted protein	Predicted protein	W4YMQ5		
AIPGENE26413	Predicted protein	Predicted protein	A7RP09		
AIPGENE26443	Predicted protein	Predicted protein	A7RP09		
AIPGENE26482	Predicted protein	Predicted protein	A7SR48		
AIPGENE26586	Predicted protein	Predicted protein	A7RLT4		
AIPGENE26659	Predicted Protein	Predicted Protein			
AIPGENE27038	Predicted protein	Predicted protein	A7RLP7		
AIPGENE27397	Predicted protein	Predicted protein	A7S3J1		
AIPGENE27454	Predicted Protein	Predicted Protein			
AIPGENE27869	Predicted protein	Predicted protein	A7SPS5		
AIPGENE28713	Predicted protein	Predicted protein	V4C0Z0		
AIPGENE3220	Predicted Protein	Predicted Protein			
AIPGENE3833	Predicted Protein	Predicted Protein			
AIPGENE3851	Predicted Protein	Predicted Protein			
AIPGENE3932	Predicted protein	Predicted protein	A7SW67		
AIPGENE4137	Predicted Protein	Predicted Protein			
AIPGENE4204	Predicted protein	Predicted protein	A7RT43	ENSCINP00000021573.3	GO:0016020 GO:0016021
AIPGENE4213	Predicted Protein	Predicted Protein	KXJ22189.1		
AIPGENE4240	Predicted protein	Predicted protein	A7RS57		
AIPGENE4275	Predicted Protein	Predicted Protein			
AIPGENE4835	Predicted Protein	Predicted Protein			
AIPGENE5210	Predicted Protein	Predicted Protein			
AIPGENE5409	Predicted protein	Predicted protein	A7SA37		
AIPGENE6480	Predicted protein	Predicted protein	A7S4R8		
AIPGENE6765	Predicted protein	Predicted protein	A7SXP8		
AIPGENE6863	Predicted Protein	Predicted Protein			
AIPGENE7096	Predicted protein	Predicted protein	V4AUI6		
AIPGENE72	Predicted Protein	Predicted Protein			
AIPGENE7613	Predicted Protein	Predicted Protein			
AIPGENE8130	Predicted Protein	Predicted Protein			
AIPGENE8511	Predicted protein	Predicted protein	I3KX76		
AIPGENE8585	Predicted protein	Predicted protein	A7S659		
AIPGENE8631	Predicted Protein	Predicted Protein			
AIPGENE8674	Predicted protein	Predicted protein	R7RXF9		
AIPGENE9919	Predicted protein	Predicted protein	A7SIY5	ENSCINP00000020132.3	GO:0005886 GO:0007411 GO:0016020 GO:0016021 GO:0046875 GO:0048013

Table S1-5. The symbiont HR-DEGs based on the annotations in *A. thaliana*

B. minutum geneID	*A. thaliana* geneID	UniProt entry name	UniProt Gene names	UniProt protein names	GOID
symbB.v1.2.001109.t1	AT1G29900.1	CARB_ARATH	CARB	Carbamoyl-phosphate synthase large chain, chloroplastic (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain) (Protein VENOSA 6)	GO:0000050 GO:0004087 GO:0004088 GO:0005524 GO:0005739 GO:0005951 GO:0006526 GO:0009507 GO:0009570 GO:0016020 GO:0016036 GO:0044205 GO:0046872
symbB.v1.2.003844.t1	AT5G63860.1	UVR8_ARATH	UVR8	Ultraviolet-B receptor UVR8 (Protein UV-B RESISTANCE 8) (RCC1 domain-containing protein UVR8)	GO:0000785 GO:0003682 GO:0005085 GO:0005515 GO:0005634 GO:0005737 GO:0005829 GO:0009411 GO:0009649 GO:0009881 GO:0010224 GO:0018298 GO:0042802 GO:0042803
symbB.v1.2.004610.t1	AT5G44510.4	A0A1P8BBP6_ARATH	TAO1	Target of AVRB operation1	GO:0005524 GO:0006952 GO:0007165 GO:0009816 GO:0042742 GO:0043531
symbB.v1.2.010303.t1	AT5G65750.1	Q9FLH2_ARATH	MPA24.10	2-oxoglutarate dehydrogenase, E1 component (Putative 2-oxoglutarate dehydrogenase E1 component)	GO:0004591 GO:0005739 GO:0008270 GO:0030976 GO:0046686 GO:0050897
symbB.v1.2.025766.t1	AT4G04740.2	F4JGW8_ARATH	CPK23	Calcium-dependent protein kinase 23	GO:0004672 GO:0004683 GO:0005509 GO:0005515 GO:0005516 GO:0005524 GO:0005634 GO:0005737 GO:0005886 GO:0006468 GO:0009738 GO:0009931 GO:0016020 GO:0016301 GO:0018105 GO:0019903 GO:0035556 GO:0046686 GO:0046777
symbB.v1.2.034993.t1	AT4G36180.1	Y4361_ARATH	At4g36180	Probable LRR receptor-like serine/threonine-protein kinase At4g36180 (EC 2.7.11.1)	GO:0004674 GO:0005524 GO:0005886 GO:0006468 GO:0007169 GO:0016021 GO:0016301
symbB.v1.2.014138.t1	AT1G47128.1	RD21A_ARATH	RD21A	Cysteine proteinase RD21A (EC 3.4.22.-) (Protein RESPONSIVE TO DEHYDRATION 21) (RD21)	GO:0000932 GO:0004197 GO:0005515 GO:0005576

					GO:0005615 GO:0005764 GO:0005773 GO:0005794 GO:0006508 GO:0008233 GO:0008234 GO:0009414 GO:0009506 GO:0009507 GO:0010494 GO:0048046 GO:0050832 GO:0051603
symbB.v1.2.004810.t1	AT1G50200.2	F4I4Z2_ARATH	ALATS	Alanine—tRNA ligase (EC 6.1.1.7) (Alanyl-tRNA synthetase) (AlaRS)	GO:0000049 GO:0004813 GO:0005524 GO:0005737 GO:0005739 GO:0005829 GO:0006400 GO:0006419 GO:0009507 GO:0016597 GO:0046686 GO:0046872
symbB.v1.2.010710.t1	AT1G61950.1	CDPKJ_ARATH	CPK19	Calcium-dependent protein kinase 19 (EC 2.7.11.1)	GO:0004683 GO:0005509 GO:0005516 GO:0005524 GO:0005634 GO:0005737 GO:0005886 GO:0006468 GO:0009738 GO:0009931 GO:0016301 GO:0018105 GO:0035556 GO:0046777
symbB.v1.2.011717.t1	AT1G64790.3	A0A1P8AT97_ARATH	ILA	ILITYHIA	GO:0005634 GO:0005829 GO:0009507 GO:0009627 GO:0016020 GO:0042742
symbB.v1.2.002493.t1	AT1G79930.1	HSP7O_ARATH	HSP70-14	Heat shock 70 kDa protein 14 (Heat shock protein 70-14) (AtHsp70-14) (Heat shock protein 91)	GO:0005515 GO:0005524 GO:0005634 GO:0005737 GO:0005829 GO:0005886 GO:0006457 GO:0009408 GO:0046686
symbB.v1.2.022102.t1	AT2G04350.2	LACS8_ARATH	LACS8	Long chain acyl-CoA synthetase 8 (EC 6.2.1.3)	GO:0000723 GO:0000784 GO:0004467 GO:0005524 GO:0005783 GO:0005794 GO:0006631 GO:0006633 GO:0009507 GO:0009941 GO:0043047 GO:0102391
symbB.v1.2.033717.t1	AT2G32400.1	GLR37_ARATH	GLR3.7	Glutamate receptor 3.7 (Ionotropic glutamate receptor GLR5) (Ligand-gated ion channel 3.7)	GO:0004970 GO:0005217 GO:0005515 GO:0005576 GO:0006874 GO:0009416 GO:0009506

					GO:0016021
symbB.v1.2.039877.t1	AT3G19170.1	PREP1_ARATH	PREP1	Presequence protease 1, chloroplastic/mitochondrial (AtPreP1) (PreP 1) (EC 3.4.24.-) (Zinc metalloprotease 1) (AtZnMP1)	GO:0004222 GO:0005515 GO:0005739 GO:0005759 GO:0006508 GO:0008233 GO:0009507 GO:0009570 GO:0009941 GO:0016485 GO:0046686 GO:0046872 GO:0048046
symbB.v1.2.003253.t1	AT3G25230.1	FKB62_ARATH	FKBP62	Peptidyl-prolyl cis-trans isomerase FKBP62 (PPIase FKBP62) (EC 5.2.1.8) (70 kDa peptidyl-prolyl isomerase) (FK506-binding protein 62) (AtFKBP62) (Immunophilin FKBP62) (Peptidylprolyl isomerase ROF1) (Protein ROTAMASE FKBP 1) (Rotamase)	GO:0003755 GO:0005515 GO:0005516 GO:0005528 GO:0005634 GO:0005737 GO:0005829 GO:0006970 GO:0009408 GO:0009611 GO:0009735 GO:0009845 GO:0032266 GO:0046686 GO:0061077 GO:0070370 GO:0071944 GO:0080025
symbB.v1.2.012494.t1	AT3G25230.1	FKB62_ARATH	FKBP62	Peptidyl-prolyl cis-trans isomerase FKBP62 (PPIase FKBP62) (EC 5.2.1.8) (70 kDa peptidyl-prolyl isomerase) (FK506-binding protein 62) (AtFKBP62) (Immunophilin FKBP62) (Peptidylprolyl isomerase ROF1) (Protein ROTAMASE FKBP 1) (Rotamase)	GO:0003755 GO:0005515 GO:0005516 GO:0005528 GO:0005634 GO:0005737 GO:0005829 GO:0006970 GO:0009408 GO:0009611 GO:0009735 GO:0009845 GO:0032266 GO:0046686 GO:0061077 GO:0070370 GO:0071944 GO:0080025
symbB.v1.2.004124.t1	AT3G58510.3	RH11_ARATH	RH11	DEAD-box ATP-dependent RNA helicase 11 (EC 3.6.4.13)	GO:0003729 GO:0005524 GO:0005730 GO:0005777 GO:0005886 GO:0008026
symbB.v1.2.018894.t1	AT4G24190.2	F4JQ55_ARATH	SHD	Chaperone protein htpG family protein	GO:0005524 GO:0005634 GO:0005739 GO:0005773 GO:0005774 GO:0005783 GO:0005788 GO:0005886 GO:0006457 GO:0009306 GO:0009409 GO:0009414 GO:0009506 GO:0009507 GO:0009651 GO:0009934 GO:0010075 GO:0016020

					GO:0016887 GO:0034976 GO:0046686 GO:0048046 GO:0051082
symbB.v1.2.028483.t1	AT4G33470.1	I14_ARATH	I14	Histone deacetylase 14 (EC 3.5.1.98)	GO:0004407 GO:0005515 GO:0005634 GO:0005829 GO:0006351 GO:0006355 GO:0009507 GO:0016575 GO:0032041 GO:0042903 GO:0043014 GO:0043621 GO:0048487 GO:0051721 GO:0090042
symbB.v1.2.006703.t1	AT5G12000.2	A0A1P8BC68_ARATH	At5g12000	Kinase with adenine nucleotide alpha hydrolases-like domain-containing protein	GO:0004672 GO:0005524 GO:0005634 GO:0006468 GO:0006950 GO:0016301 GO:0016787
symbB.v1.2.015655.t1	AT5G14640.2	KSG5_ARATH	ASK5	Shaggy-related protein kinase epsilon (EC 2.7.11.1) (ASK-epsilon) (Shaggy-related protein kinase 13) (AtSK13)	GO:0004672 GO:0004674 GO:0005515 GO:0005524 GO:0005829 GO:0006972 GO:0009507 GO:0009651 GO:0046777
symbB.v1.2.040764.t1	AT5G15450.1	CLPB3_ARATH	CLPB3	Chaperone protein ClpB3, chloroplastic (ATP-dependent Clp protease ATP-binding subunit ClpB homolog 3) (Casein lytic proteinase B3) (Protein ALBINO OR PALE GREEN 6)	GO:0005524 GO:0005737 GO:0009408 GO:0009507 GO:0009532 GO:0009570 GO:0009658 GO:0016485 GO:0016887
symbB.v1.2.001344.t1	AT5G22060.1	DNAJ2_ARATH	ATJ2	Chaperone protein dnaJ 2 (AtDjA2)	GO:0005524 GO:0005737 GO:0005829 GO:0005886 GO:0006457 GO:0009408 GO:0009506 GO:0031072 GO:0046872 GO:0051082
symbB.v1.2.040261.t1	AT5G23860.2	TBB8_ARATH	TUBB8	Tubulin beta-8 chain (Beta-8-tubulin)	GO:0003924 GO:0005200 GO:0005515 GO:0005525 GO:0005737 GO:0005794 GO:0005874 GO:0007017 GO:0009409 GO:0009651 GO:0010583 GO:0016020 GO:0045298
symbB.v1.2.040262.t1	AT5G23860.2	TBB8_ARATH	TUBB8	Tubulin beta-8 chain (Beta-8-tubulin)	GO:0003924 GO:0005200 GO:0005515 GO:0005525 GO:0005737 GO:0005794

					GO:0005874 GO:0007017 GO:0009409 GO:0009651 GO:0010583 GO:0016020 GO:0045298
symbB.v1.2.036077.t1	AT5G26860.1	LONM1_ARATH	LON1	Lon protease homolog 1, mitochondrial (EC 3.4.21.53)	GO:0003677 GO:0004176 GO:0004252 GO:0005515 GO:0005524 GO:0005739 GO:0005759 GO:0005829 GO:0006515 GO:0007005 GO:0008236
symbB.v1.2.018019.t1	AT5G38480.1	14333_ARATH	GRF3	14-3-3-like protein GF14 psi (General regulatory factor 3) (Protein RARE COLD INDUCIBLE 1A)	GO:0005515 GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005739 GO:0005773 GO:0005794 GO:0005829 GO:0005886 GO:0006995 GO:0009409 GO:0009506 GO:0009507 GO:0009631 GO:0009873 GO:0016036 GO:0019222 GO:0019904 GO:0045309 GO:0050826 GO:0051365
symbB.v1.2.026274.t1	AT5G38480.1	14333_ARATH	GRF3	14-3-3-like protein GF14 psi (General regulatory factor 3) (Protein RARE COLD INDUCIBLE 1A)	GO:0005515 GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005739 GO:0005773 GO:0005794 GO:0005829 GO:0005886 GO:0006995 GO:0009409 GO:0009506 GO:0009507 GO:0009631 GO:0009873 GO:0016036 GO:0019222 GO:0019904 GO:0045309 GO:0050826 GO:0051365
symbB.v1.2.006004.t1	AT5G42020.1	MD37F_ARATH	MED37F	Mediator of RNA polymerase II transcription subunit 37f (Heat shock 70 kDa protein 12) (Heat shock protein 70-12) (AtHsp70-12) (Luminal-binding protein 2) (AtBP2) (BiP2)	GO:0005524 GO:0005618 GO:0005634 GO:0005730 GO:0005773 GO:0005774 GO:0005783 GO:0005788 GO:0005886 GO:0006351 GO:0006355 GO:0006457

					GO:0009408 GO:0009506 GO:0009507 GO:0009735 GO:0010197 GO:0016020 GO:0016592 GO:0034976 GO:0046686
symbB.v1.2.028955.t1	AT5G50920.1	CLPC1_ARATH	CLPC1	Chaperone protein ClpC1, chloroplastic (ATP-dependent Clp protease ATP-binding subunit ClpC homolog 1) (Casein lytic proteinase C1) (Protein DE-REGULATED CAO ACCUMULATION 1) (Protein IRON-RESCUED MUTANT 1)	GO:0004176 GO:0005515 GO:0005524 GO:0005618 GO:0005739 GO:0009507 GO:0009532 GO:0009535 GO:0009536 GO:0009570 GO:0009658 GO:0009706 GO:0009735 GO:0009941 GO:0010380 GO:0016887 GO:0031897 GO:0045036 GO:0045037
symbB.v1.2.030369.t1	AT5G52640.1	HS901_ARATH	HSP90-1	Heat shock protein 90-1 (AtHSP90.1) (AtHsp90-1) (Heat shock protein 81-1) (Hsp81-1) (Heat shock protein 83)	GO:0005515 GO:0005524 GO:0005618 GO:0005737 GO:0005829 GO:0005886 GO:0009408 GO:0009816 GO:0042742 GO:0046685 GO:0051082 GO:0061077
symbB.v1.2.015892.t1	AT5G56000.1	HS904_ARATH	HSP90-4	Heat shock protein 90-4 (AtHSP90.4) (AtHsp90-4) (Heat shock protein 81-4) (Hsp81-4)	GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005774 GO:0005794 GO:0005829 GO:0005886 GO:0006457 GO:0006950 GO:0009570 GO:0048046 GO:0051082
symbB.v1.2.029700.t1	AT5G56000.1	HS904_ARATH	HSP90-4	Heat shock protein 90-4 (AtHSP90.4) (AtHsp90-4) (Heat shock protein 81-4) (Hsp81-4)	GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005774 GO:0005794 GO:0005829 GO:0005886 GO:0006457 GO:0006950 GO:0009570 GO:0048046 GO:0051082
symbB.v1.2.029704.t1	AT5G56000.1	HS904_ARATH	HSP90-4	Heat shock protein 90-4 (AtHSP90.4) (AtHsp90-4) (Heat shock protein 81-4) (Hsp81-4)	GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005774 GO:0005794 GO:0005829 GO:0005886

					GO:0006457 GO:0006950 GO:0009570 GO:0048046 GO:0051082
symbB.v1.2.018945.t1	AT5G56030.1	HS902_ARATH	HSP90-2	Heat shock protein 90-2 (AtHSP90.2) (AtHsp90-2) (Heat shock protein 81-2) (Hsp81-2) (Protein EARLY-RESPONSIVE TO DEHYDRATION 8) (Protein LOSS OF RECOGNITION OF AVRRPM1 2) (Protein MUTANT SNC1-ENHANCING 12)	GO:0003729 GO:0005515 GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005739 GO:0005794 GO:0005829 GO:0006457 GO:0006952 GO:0009408 GO:0009414 GO:0009506 GO:0009651 GO:0009816 GO:0009908 GO:0010187 GO:0010286 GO:0016887 GO:0048366 GO:0050821 GO:0051082 GO:0051131 GO:0061077 GO:0090332
symbB.v1.2.033132.t1	AT5G56030.1	HS902_ARATH	HSP90-2	Heat shock protein 90-2 (AtHSP90.2) (AtHsp90-2) (Heat shock protein 81-2) (Hsp81-2) (Protein EARLY-RESPONSIVE TO DEHYDRATION 8) (Protein LOSS OF RECOGNITION OF AVRRPM1 2) (Protein MUTANT SNC1-ENHANCING 12)	GO:0003729 GO:0005515 GO:0005524 GO:0005618 GO:0005634 GO:0005737 GO:0005739 GO:0005794 GO:0005829 GO:0006457 GO:0006952 GO:0009408 GO:0009414 GO:0009506 GO:0009651 GO:0009816 GO:0009908 GO:0010187 GO:0010286 GO:0016887 GO:0048366 GO:0050821 GO:0051082 GO:0051131 GO:0061077 GO:0090332
symbB.v1.2.029800.t1	AT5G56290.1	PEX5_ARATH	PEX5	Peroxisome biogenesis protein 5 (Peroxin-5) (AtPEX5) (Peroxisomal targeting signal type 1 receptor) (Pex5p)	GO:0005052 GO:0005515 GO:0005634 GO:0005778 GO:0005829 GO:0006625 GO:0009733 GO:0016560
symbB.v1.2.010467.t1	AT5G58140.4	F4KDJ3_ARATH	PHOT2	Phototropin 2	GO:0000155 GO:0000160 GO:0004674 GO:0005515 GO:0005524 GO:0005634 GO:0005794 GO:0005886 GO:0007623

					GO:0009637 GO:0009638 GO:0009735 GO:0009882 GO:0009902 GO:0010118 GO:0010181 GO:0010362 GO:0016020 GO:0016301 GO:0018298 GO:0035556 GO:0042802 GO:0046777
symbB.v1.2.006972.t1	AT5G66120.1	DHQS_ARATH	DHQS	3-dehydroquinate synthase, chloroplastic (EC 4.2.3.4)	GO:0003856 GO:0005737 GO:0005768 GO:0005794 GO:0005802 GO:0009073 GO:0009423 GO:0009507 GO:0009570 GO:0046872
symbB.v1.2.000094.t1	AT1G35190.1	Q9C6F0_ARATH	At1g35190	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (At1g35190) (Hyoscyamine 6-dioxygenase hydroxylase, putative)	GO:0005737 GO:0009821 GO:0016491 GO:0016706 GO:0046872 GO:0051213 GO:0055114
symbB.v1.2.000744.t1	AT2G25070.2	P2C21_ARATH	PPC4-2	Probable protein phosphatase 2C 21 (AtPP2C21) (EC 3.1.3.16) (AtPPC4;2)	GO:0004722 GO:0005634 GO:0005886 GO:0006470 GO:0009737 GO:0046872
symbB.v1.2.005378.t1	AT4G10120.4	A0A1P8B6E2_ARATH	ATSPS4F	Sucrose-phosphate synthase family protein	GO:0005634 GO:0005886 GO:0016757 GO:0046524
symbB.v1.2.006657.t1	AT3G04710.2	F4J4V4_ARATH	TPR10	Ankyrin repeat family protein	GO:0005575 GO:0008150
symbB.v1.2.009347.t1	AT5G64300.1	RIBA1_ARATH	RIBA1	Bifunctional riboflavin biosynthesis protein RIBA 1, chloroplastic (AtRIBA1) [Includes: 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBP synthase) (EC 4.1.99.12); GTP cyclohydrolase-2 (EC 3.5.4.25) (GTP cyclohydrolase II)]	GO:0003935 GO:0005525 GO:0008686 GO:0009231 GO:0009507 GO:0009570 GO:0016020 GO:0046872
symbB.v1.2.009837.t1	AT2G47840.1	TI202_ARATH	TIC20-II	Protein TIC 20-II, chloroplastic (Translocon at the inner envelope membrane of chloroplasts 20-II) (AtTIC20-II)	GO:0005739 GO:0009507 GO:0009536 GO:0009706 GO:0009941 GO:0015031 GO:0016021
symbB.v1.2.012069.t1	AT4G34570.5	A0A1P8B676_ARATH	THY-2	Bifunctional dihydrofolate reductase-thymidylate synthase	GO:0004146 GO:0004799 GO:0006231 GO:0006545 GO:0006730 GO:0009165 GO:0009257 GO:0009507 GO:0032259 GO:0046654 GO:0050661 GO:0055114
symbB.v1.2.012814.t1	AT5G42620.2	F4K306_ARATH	MFO20.3	Metalloendopeptidase / zinc ion binding protein	GO:0005886
symbB.v1.2.015277.t1	AT3G59260.1	PRNL4_ARATH	At3g59260	Putative pirin-like protein At3g59260	GO:0005516 GO:0005634

symbB.v1.2.015898.t1	AT5G45400.1	RFA1C_ARATH	RPA1C	Replication protein A 70 kDa DNA-binding subunit C (AtRPA70c) (AtRPA1-2) (Replication factor A protein 1C) (Replication protein A 1C) (AtRPA1C)	GO:0003677 GO:0005634 GO:0006260 GO:0006281 GO:0006310 GO:0046872
symbB.v1.2.016011.t1	AT3G24300.1	AMT13_ARATH	AMT1-3	Ammonium transporter 1 member 3 (AtAMT1;3)	GO:0005886 GO:0005887 GO:0006810 GO:0008519 GO:0010311 GO:0015695 GO:0015696 GO:0016020 GO:0072488 GO:0080181
symbB.v1.2.022050.t1	AT1G76340.1	GONS3_ARATH	GONST3	GDP-mannose transporter GONST3 (Protein GOLGI NUCLEOTIDE SUGAR TRANSPORTER 3)	GO:0000139 GO:0005886 GO:0008643 GO:0015297 GO:0015780 GO:0016020 GO:0016021
symbB.v1.2.023017.t1	AT3G06460.1	Q9SQU9_ARATH	At3g06460	F24P17.4 protein (GNS1/SUR4 membrane protein family)	GO:0003674 GO:0008150 GO:0016021
symbB.v1.2.023478.t1	AT1G03070.3	LFG4_ARATH	LFG4	Protein LIFEGUARD 4 (AtLFG4)	GO:0008150 GO:0016021 GO:0016595
symbB.v1.2.025087.t1	AT1G11390.1	F4I8V6_ARATH	At1g11390	Protein kinase superfamily protein	GO:0005886
symbB.v1.2.029381.t1	AT2G16960.2	Q1PF58_ARATH	At2g16960	ARM repeat superfamily protein (Importin beta-2 subunit family protein)	GO:0005634 GO:0008565 GO:0015031
symbB.v1.2.030912.t1	AT2G40770.3	F4II36_ARATH	At2g40770	RING-finger, DEAD-like helicase, PHD and SNF2 domain-containing protein	GO:0005634
symbB.v1.2.033324.t1	AT2G32390.5	GLR35_ARATH	GLR3.5	Glutamate receptor 3.5 (Ionotropic glutamate receptor GLR6) (Ligand-gated ion channel 3.5)	GO:0004930 GO:0004970 GO:0005216 GO:0005217 GO:0005234 GO:0005739 GO:0005886 GO:0006811 GO:0006874 GO:0007186 GO:0009416 GO:0009507 GO:0009845 GO:0016021 GO:0070584
symbB.v1.2.035092.t1	AT5G55760.1	SIR1_ARATH	SRT1	NAD-dependent protein deacetylase SRT1 (EC 3.5.1.-) (Regulatory protein SIR2 homolog 1) (SIR2-like protein 1)	GO:0003677 GO:0005634 GO:0005677 GO:0005737 GO:0006342 GO:0006355 GO:0016787 GO:0046872 GO:0070403
symbB.v1.2.035131.t1	AT1G24610.1	Q9FYK3_ARATH	At1g24610	F21J9.27 (Rubisco methyltransferase family protein)	GO:0009507
symbB.v1.2.014119.t1	AT3G29390.2	RIK_ARATH	RIK	Protein RIK (Rough sheath 2-interacting KH domain protein) (RS2-interacting KH domain protein)	GO:0003723 GO:0005634 GO:0008150
symbB.v1.2.000553.t1					
symbB.v1.2.000729.t2					
symbB.v1.2.002453.t1					
symbB.v1.2.004489.t1					
symbB.v1.2.006190.t1					
symbB.v1.2.006485.t1					
symbB.v1.2.006662.t1					

symbB.v1.2.006674.t1
symbB.v1.2.006939.t1
symbB.v1.2.007805.t1
symbB.v1.2.009050.t1
symbB.v1.2.009338.t2
symbB.v1.2.009845.t1
symbB.v1.2.010238.t1
symbB.v1.2.011248.t1
symbB.v1.2.012438.t1
symbB.v1.2.012498.t1
symbB.v1.2.012792.t1
symbB.v1.2.014506.t1
symbB.v1.2.014634.t1
symbB.v1.2.015952.t1
symbB.v1.2.017503.t1
symbB.v1.2.018823.t1
symbB.v1.2.019579.t1
symbB.v1.2.019791.t1
symbB.v1.2.020598.t1
symbB.v1.2.021635.t1
symbB.v1.2.022234.t1
symbB.v1.2.022714.t1
symbB.v1.2.023175.t1
symbB.v1.2.023346.t1
symbB.v1.2.024748.t1
symbB.v1.2.024757.t1
symbB.v1.2.024924.t1
symbB.v1.2.025040.t1
symbB.v1.2.025131.t1
symbB.v1.2.025135.t1
symbB.v1.2.025283.t1
symbB.v1.2.025497.t1
symbB.v1.2.026162.t1
symbB.v1.2.027630.t1
symbB.v1.2.028077.t1
symbB.v1.2.028225.t1
symbB.v1.2.028298.t1
symbB.v1.2.028741.t1
symbB.v1.2.029913.t1
symbB.v1.2.030149.t1
symbB.v1.2.031586.t1
symbB.v1.2.031823.t1
symbB.v1.2.032324.t1
symbB.v1.2.033518.t1
symbB.v1.2.033681.t1
symbB.v1.2.035210.t1
symbB.v1.2.035467.t1
symbB.v1.2.035741.t1
symbB.v1.2.036226.t1
symbB.v1.2.036945.t1
symbB.v1.2.037278.t1
symbB.v1.2.037360.t1
symbB.v1.2.038399.t1
symbB.v1.2.039367.t1
symbB.v1.2.040571.t1
symbB.v1.2.040844.t1
symbB.v1.2.041086.t1

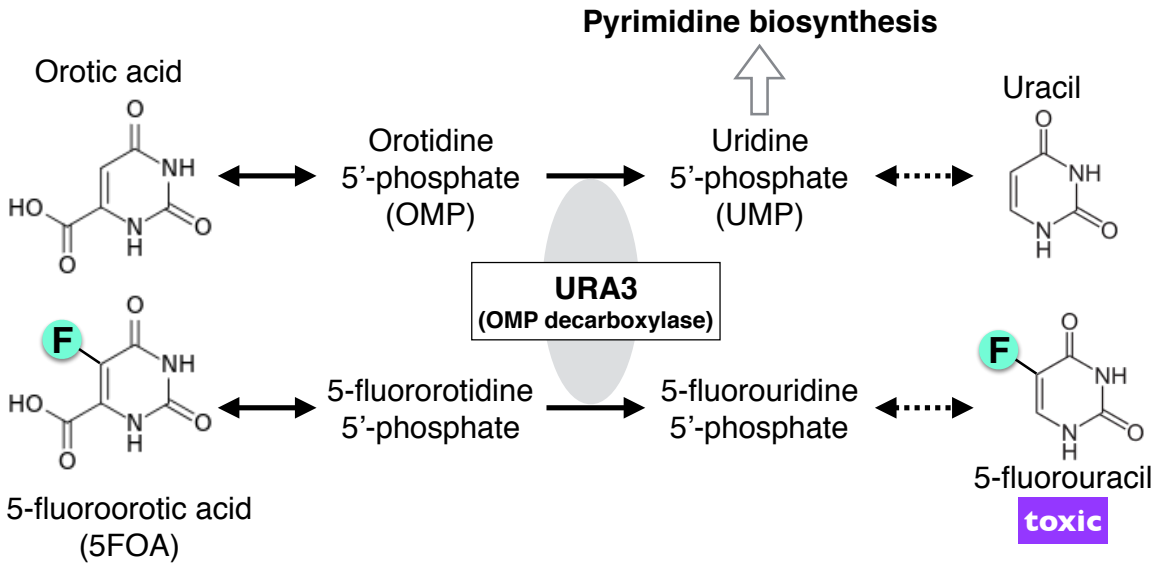
Supplementary tables and figures 2

Table S2-1. Primers used for sequencing

Primer name	Seq (5' to 3')	Start (scaffold311.1)
Ura3c_F1	TATGTGTGGGCCTTGATCCAC	122845
Ura3c_R1	ATCATTCTCATCTCGCCGCTC	127568
Ura3c_R2	TCCATGGCATCTACATCGGTTG	126328
Ura3g_F1	ATCACATCCCACCTCATGAACC	121308
Ura3g_F2	CATCACTGCAAAACTGATAAGGGCC	121968
Ura3g_F3	GAAGGGTTTTGGTGAAGGTTGAGGG	122434
Ura3g_F4	TTTTGGGTGCATCATAACGTCATGG	123246
Ura3g_F5	TGAACTTTACCAGCGCCCATTTT	123413
Ura3g_F6	TGCGAGTCTCTTCATTTTGGCAA	124468
Ura3g_F7	CATGGCTGTATGATATGAGCAACGC	124790
Ura3g_F8	TTTCATGGTGACTGGGCTGATAAGA	126683
Ura3g_F9	TTGGTTGATTCTCCCCACCA	126600
Ura3g_R1	CCCAGTTGTGTCATGCTCATTC	127724
Ura3g_R2	TGCTCATTCATTTACACCTGTCCTGA	127711
Ura3g_R3	ACCAAGACTAACAGTACGCACCT	126002
Ura3g_R4	CTTTAATGACAATGGCCCGGCTA	125045
Ura3g_R5	CCACCTCCACAGAAAGAACATGC	124041
Ura3g_R6	AAGTTCGTTGTAAAGCTTCCCAACC	123018
Ura3g_R7	TGGATTCAATCATTTCACCAACCCA	125583
Ura3g_R8	TCCAACCTCTAGATATTGGCAGCAA	127141
Ura3g_R9	GGTTAATAAGCCCGGCCTTG	121729
Ura3g_R10	CTAGTGAGGAGTTTTGGCGC	122024
Ura3g_R11	TCAAACAAGACATGCCAGCC	122141
Ura3g_R12	TCCTCAATGAGATGCTGGCA	122282
T7	TAATACGACTCACTATAGGG	
SP6m	CTCCCATATGGTCGACCTGC	

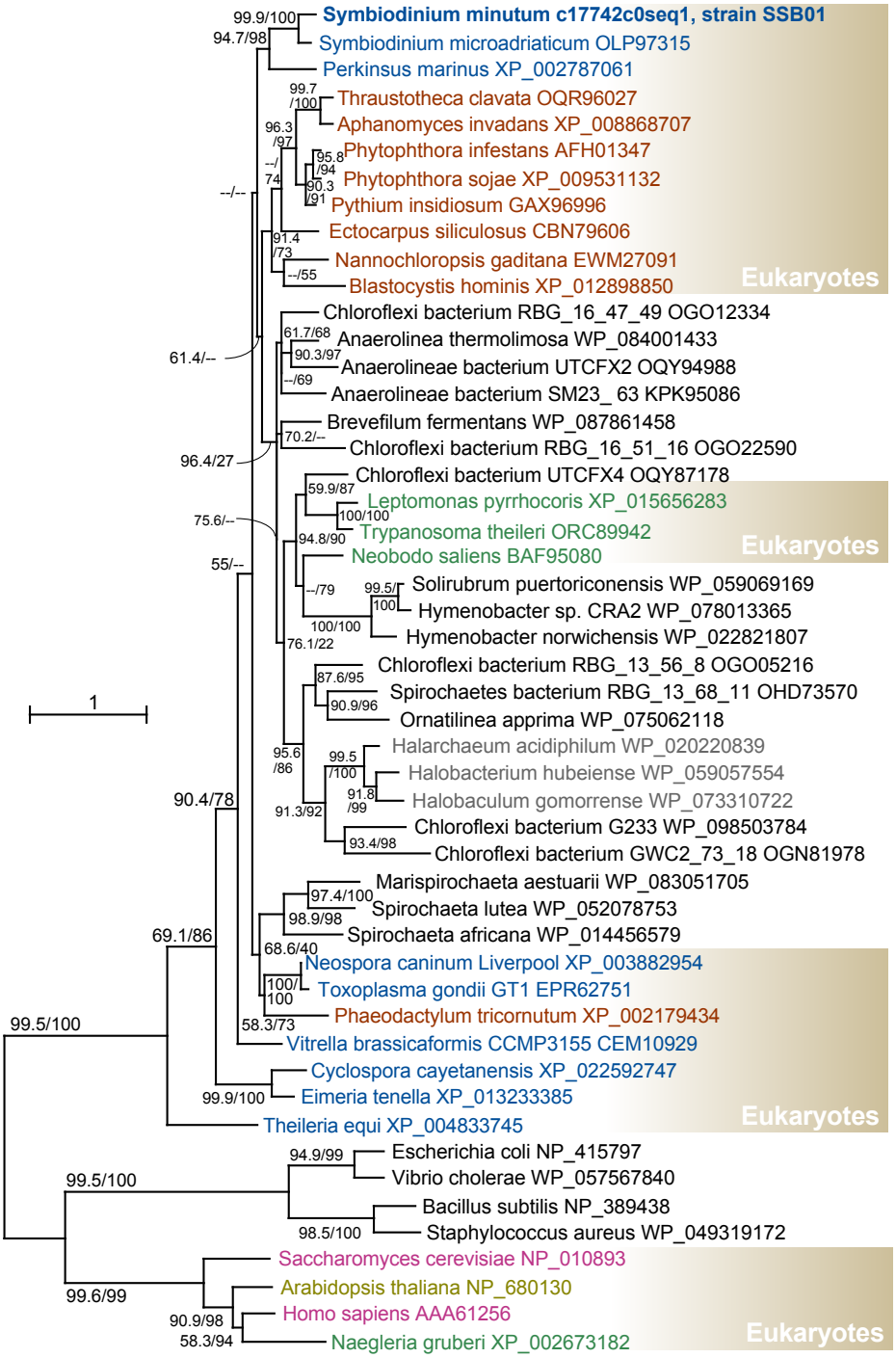
Supplementary Fig. S2-1**. Reactions catalysed by OMP decarboxylase, the *URA3* gene product.**

Cells possessing active *URA3* that encodes OMP decarboxylase can convert OMP to UMP, and are sensitive to 5FOA due to the lethal toxicity of 5-fuorouracil, while mutants possessing non-functional *URA3* gene are unable to synthesize pyrimidine derivatives, but are resistant to 5FOA.



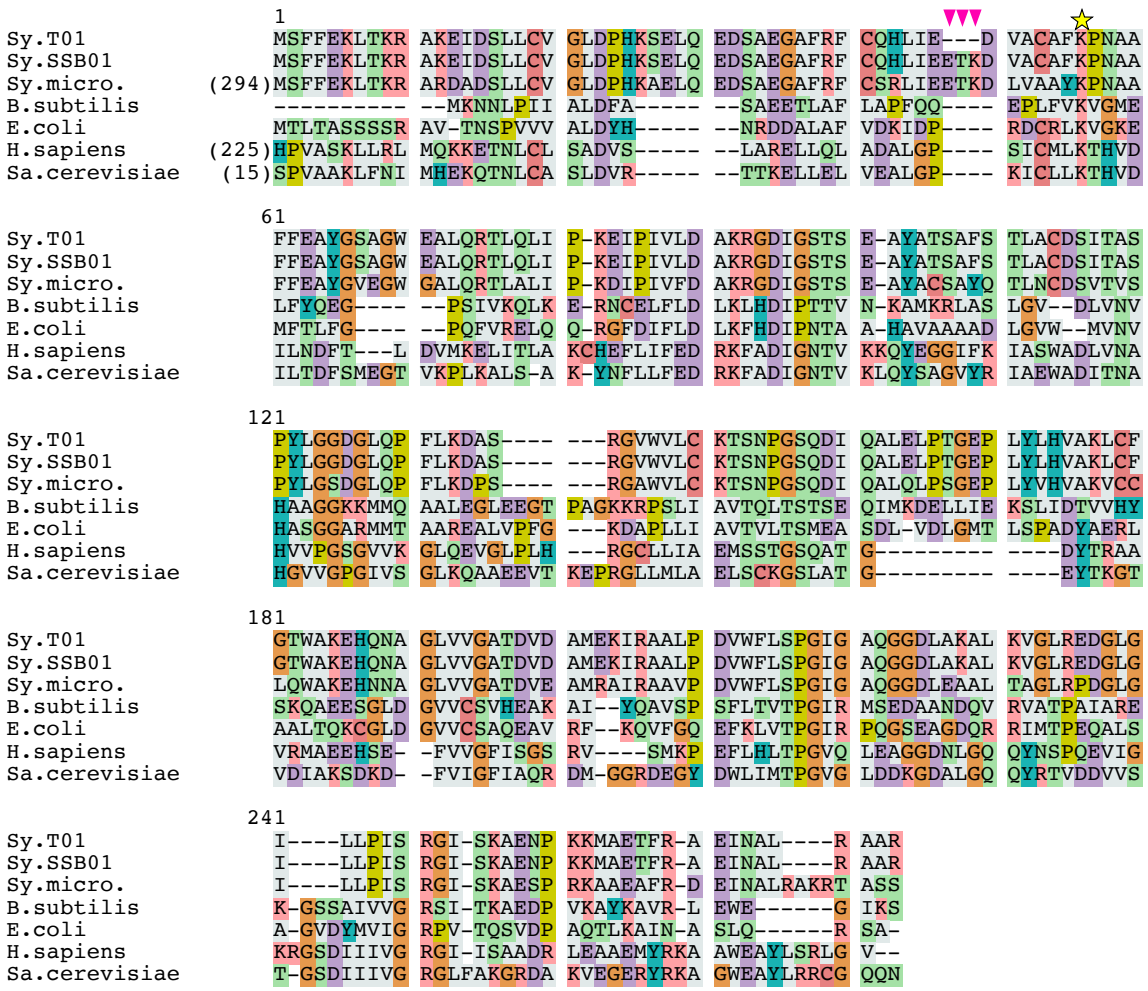
Supplementary Fig. S2-2. **Sequence analysis of URA3 proteins**

A. Phylogenetic tree of URA3 proteins. Support values using SH-like approximate likelihood ratio test (left) and ultrafast bootstrap approximation (right) are shown on each branch. Text colors represent phylogenetic domains or eukaryotic super-groups to which the organisms belong. Black, Bacteria; gray, Archaea; blue, Alveolata; brown, Stramenopiles; green, Excavata; olive, Archaeplastida; magenta, Opisthokonta. A single copy *URA3* homolog was found in *Symbiodinium kawagutii* (Skav218766) but not used due to its highly divergent and non-alignable sequence.



Supplementary Fig. S2-2. **Sequence analysis of URA3 proteins** (continued)

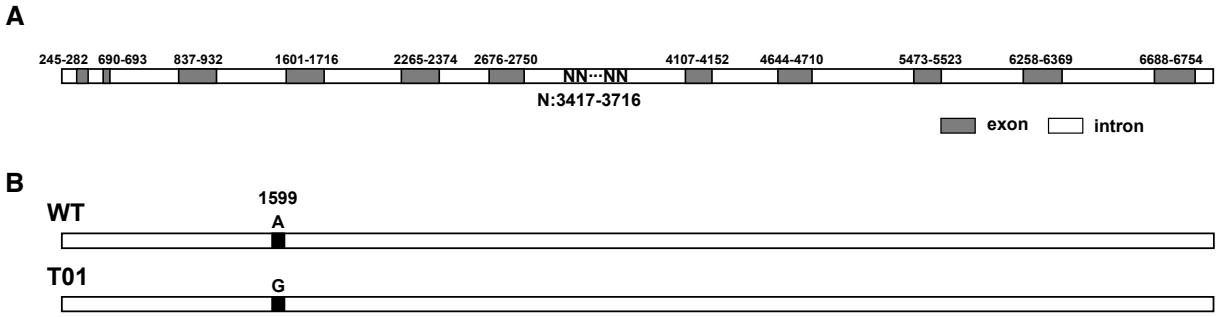
B. Multiple amino acid sequence alignment of URA3 proteins. Star and triangle indicate a conserved lysine residue and the deletion site found in T01, respectively. Proteins from *Symbiodinium* sp. T01 (sequenced in this study), *Symbiodinium* sp. SSB01 (identical to *Sy. minutum* contig comp17742_c0_seq1), *Sy. microadriaticum* (accession number: OLP97315), *Bacillus subtilis* (NP_389438), *Escherichia coli* (NP_415797), *Homo sapiens* (AAA61256), *Saccharomyces cerevisiae* (NP_010893) were used to construct the alignment.



Supplementary Fig. S2-3. **Genomic sequences of *URA3* gene in wild type and T01**

A. WT *URA3* gene structure based on the gene model of *S. minutum* Mf1.05b is shown.

B. In T01, the 1599th base has been substituted from A to G.



Supplementary Fig. S2-4. **Effect of uracil on the symbiotic states of wild type and T01**

To examine the effect of uracil on the symbiotic states of SSB01 (WT) and T01, *Symbiodinium* area ratios were quantified by comparing signal regions of chlorophyll autofluorescence and normalized to the value of day 0 in the presence (T01, closed circles) or absence (T01, open circles; WT, open square) of uracil.